

NASA
CP
2138-
pt.1
c.1

NASA Conference Publication 2138

LOAN COPY: RET
AFWL TECH. DIV.
KIRTLAND AFB, TX

0067349



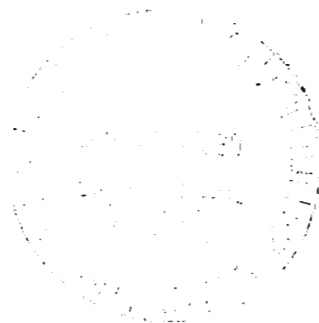
TECH LIBRARY KAFB, NM

Heterodyne Systems and Technology

Part I

Proceedings of an
international conference
held at Williamsburg, Virginia
March 25-27, 1980

NASA





NASA Conference Publication 2138

Heterodyne Systems and Technology

Part I

**Proceedings of an
international conference
held at Williamsburg, Virginia
March 25-27, 1980**



**National Aeronautics
and Space Administration**

**Scientific and Technical
Information Branch**

1980



PREFACE

The proceedings of the International Conference on Heterodyne Systems and Technology held in Williamsburg, Virginia March 25-27, 1980, are reported in this NASA Conference Publication. This conference was sponsored by the National Aeronautics and Space Administration.

The conference was the first of its kind, bringing together technologists, systems engineers, and applications scientists to exchange technical information on all aspects of optical heterodyning. Topics covered various aspects of heterodyning throughout the electromagnetic spectrum including detectors, local oscillators, tunable diode lasers, astronomical systems, and environmental sensors, with both active and passive systems represented.

The conference organization consisted of a General Program Chairman, Robert H. Kingston of M.I.T. Lincoln Laboratory, and a Program Committee whose members are listed for reference elsewhere in these proceedings. The conference organizers were Stephen J. Katzberg and James M. Hoell; the conference coordinator was Patricia Hurt. Conference support was led by Jane T. Everett with assistance from Vivian Parrous and Sheila A. Armstrong, all from Langley Research Center.

Use of trade names or names of manufacturers in this report does not constitute an official endorsement of such products or manufacturers, either expressed or implied, by NASA.

Stephen J. Katzberg
James M. Hoell

Conference Cochairmen



FOREWORD

Robert H. Kingston
Lincoln Laboratory, Massachusetts Institute of Technology
Lexington, Massachusetts 02173

heterodyne - (hetero + dyne) - separate force

superheterodyne - (super(sonic) + heterodyne)

To introduce the proceedings of this conference, it is interesting to review the origin of the above two words and trace the early history of heterodyne systems up to the first optical heterodyne experiment of 1955. The story begins in 1902 with the issue of a patent to Fessenden with the simple title, Wireless Signaling (1). The invention was the concept of transmitting two separate radio frequencies, receiving the same on two separate antennas and mixing the two to produce an audio frequency. The "mixer" in this case was an iron-core coil driving a telephone diaphragm, and the deflection was proportional to electrical power, thus yielding acoustic output at the difference frequency. A curious and ironic phrase in the patent reads: "By the term 'electromagnetic waves' as used herein is meant electromagnetic waves long in period compared with that are called 'heat-waves' or 'radiant heat'."

It is not clear when the term "heterodyne" was coined, but we find a paper by Hogan (2) in 1913 with the title "The heterodyne receiving system, and notes on the recent Arlington-Salem tests." One section of this paper is particularly significant as it reads, "It is evident that an economy may be brought about by transmitting only one wave and generating the second frequency at the receiving station." And thus came the birth of the local oscillator; "The heterodyne...is seen to consist of a standard receiving set associated with a local generating circuit by means of an inductive coupler. The generator G may be an alternator...which is a 2 K.W. machine capable of generating frequencies up to 100,000 cycles per second." This is probably a record value for local oscillator power, radio or optical. Unfortunately, this first heterodyne system was adequate for c.w. or Morse code signals, but seriously distorted voice communications.

The next step in the development of the heterodyne technique was taken by Armstrong, later the inventor and developer of f.m. broadcasting. During World War I, as a member of the U.S. Army Signal Corps, he devised a technique for receiving and separating signals over a wide range of radio frequencies using a modification of the heterodyne principle. To quote from his 1921 paper (3): "This expedient consists in reducing the frequency of the incoming signal to some predetermined super-audible frequency which can be readily amplified, passing this current through an amplifier, and then detecting or rectifying the amplified current." Thus came the superheterodyne, supersonic being the pre-jet age version of ultrasonic. In a technical discussion still valid today, Armstrong points out the difficulties of r.f. amplification (poor vacuum tube amplifiers) and the difficulties of audio amplification (poor local oscillator stability and high tube noise at audio frequencies). Armstrong's invention was

in response to the problem which he proposed at the beginning of his paper: "The problem may be summed up in the following words — to construct a receiver for undamped, modulated continuous, and damped oscillations which is substantially equally sensitive over a range of wavelengths from 50 to 600 meters, which is capable of rapid adjustment from one wave to another, and which does not distort or lose any characteristic note or tone inherent in the transmitter." By this time, the mixers were diode rectifiers or vacuum tubes and, since the early thirties, the superheterodyne has been the standard method of radio wave reception.

The next two steps which led to the development of heterodyne systems from the sub-millimeter to the visible occurred in 1955 and 1960. First, Forrester, Gudmundsen, and Johnson (4) demonstrated photoelectric mixing of the Zeeman components of a mercury emission line thus establishing the feasibility of optical heterodyning. Then, in 1960, Maiman (5) gave the first demonstration of optical maser or laser action, thus giving promise of a local oscillator and mixer which in Armstrong's words could "reduce the frequency of the incoming signal to some predetermined super-audible (!) frequency which can be readily amplified."

The papers in these proceedings take up the story from there. The applications are of course much broader than simple signal reception and include spectroscopy, radiometry, communications, and radar. The critical technologies are the design and fabrication of local oscillators and detectors. We are all indebted to the staff of NASA Langley Research Center for initiating and organizing the conference and for the preparation of these proceedings, the first such work devoted exclusively to the exciting field of heterodyne systems at really short wavelengths.

REFERENCES

1. U. S. Patent No., 706, 740, August 12, 1902.
2. J. L. Hogan, Jr., Proc. I.R.E., 1, 75, (1913).
3. E. H. Armstrong, Proc. I.R.E., 9, 3, (1921).
4. Forrester, Gudmundsen, and Johnson, Phys. Rev., 99, 1691, (1955).
5. T. H. Maiman, Nature, 187, 493, (1960).

CONTENTS

Part 1

PREFACE	iii
FOREWORD	v

SESSION I - PLENARY SESSION Chairman: Robert Menzies

1. THE PHYSICS OF HETERODYNE DETECTION IN THE FAR-INFRARED: TRANSITION FROM ELECTRIC-FIELD TO PHOTON-ABSORPTION DETECTION IN A SIMPLE SYSTEM	1
Malvin Carl Teich	
2. INFRARED HETERODYNE SPECTROSCOPY IN ASTRONOMY	11
Albert Betz	

SESSION II - LOCAL OSCILLATORS A Chairman: James Hutchby

3. NEW PbSnTe HETEROJUNCTION LASER DIODE STRUCTURES WITH IMPROVED PERFORMANCE	23
C. G. Fonstad, D. Kasemset, H. H. Hsieh, and S. Rotter	
4. ADVANCES IN TUNABLE DIODE LASER TECHNOLOGY	33
Wayne Lo	
5. LONG WAVELENGTH PbSnTe LASERS WITH CW OPERATION ABOVE 77K	45
Koji Shinohara, Mitsuo Yoshikawa, Michiharu Ito, and Ryuiti Ueda	
6. WIDELY TUNABLE (PbSn)Te LASERS USING ETCHED CAVITIES FOR MASS PRODUCTION	55
Matthew D. Miller	
7. RELIABILITY IMPROVEMENTS IN TUNABLE $Pb_{1-x}Sn_xSe$ DIODE LASERS OPERATING IN THE 8.5-30 MICROMETER SPECTRAL REGION	63
Kurt J. Linden, Jack F. Butler, Kenneth W. Nill, and Robert E. Reeder	
8. LONG WAVELENGTH $Pb_{1-x}Sn_xSe$ AND $Pb_{1-x}Sn_xTe$ DIODE LASERS AS LOCAL OSCILLATORS	79
E. R. Washwell	

9. DIAGNOSTICS AND CONTROL OF WAVENUMBER STABILITY AND PURITY OF TUNABLE DIODE LASERS RELEVANT TO THEIR USE AS LOCAL OSCILLATORS IN HETERODYNE SYSTEMS	97
S. Poultney, D. Chen, G. Steinberg, F. Wu, A. Pires, M. Miller, and M. McNally	
10. EXCESS NOISE IN $Pb_{1-x}Sn_xSe$ SEMICONDUCTOR LASERS	129
Charles N. Harward and Barry D. Sidney	
11. FINE WAVELENGTH ID FOR TUNABLE LASER LOCAL OSCILLATORS	143
M. G. Savage and R. C. Augeri	
12. A REVIEW OF THE NASA/OAST CRYOGENIC COOLERS TECHNOLOGY PROGRAM . . .	153
J. G. Lundholm, Jr. and Allan Sherman	

SESSION III - INCOHERENT APPLICATIONS
Chairman: Albert Betz

13. ATMOSPHERIC STUDIES AND APPLICATIONS WITH INFRARED HETERODYNE DETECTION TECHNIQUES	171
Robert T. Menzies	
14. INFRARED HETERODYNE RADIOMETER FOR AIRBORNE ATMOSPHERIC TRANSMITTANCE MEASUREMENTS	181
J. M. Wolczok, R. A. Lange, and A. J. DiNardo	
15. DEVELOPMENT AND PERFORMANCE OF A LASER HETERODYNE SPECTROMETER USING TUNEABLE SEMICONDUCTOR LASERS AS LOCAL OSCILLATORS	199
D. Glenar, T. Kostiuik, D. E. Jennings, and M. J. Mumma	
16. ATMOSPHERIC SOLAR ABSORPTION MEASUREMENTS IN THE 9-11 MICRON REGION USING A DIODE LASER HETERODYNE SPECTROMETER	209
Charles N. Harward and James M. Hoell, Jr.	
17. SENSITIVITY STUDIES AND LABORATORY MEASUREMENTS FOR THE LASER HETERODYNE SPECTROMETER EXPERIMENT	221
Frank Allario, Stephen J. Katzberg, and Jack C. Larsen	
18. A 163 MICRON LASER HETERODYNE RADIOMETER FOR OH: PROGRESS REPORT	241
H. M. Pickett and T. L. Boyd	
19. HETERODYNE DETECTION OF THE 752.033-GHz H_2O ROTATIONAL ABSORPTION LINE	249
G. F. Dionne, J. F. Fitzgerald, T-S. Chang, M. M. Litvak, and H. R. Fetterman	

Part 2*

SESSION IV - DETECTORS
Chairman: David Spears

20. n-p (Hg,Cd)Te PHOTODIODES FOR 8-14 MICROMETER HETERODYNE APPLICATIONS	263
J. F. Shanley and C. T. Flanagan	
21. CONCEPTUAL DESIGN AND APPLICATIONS OF HgCdTe INFRARED PHOTODIODES FOR HETERODYNE SYSTEMS	281
Michel B. Sirieix and Henry Hofheimer	
22. COMPARATIVE PERFORMANCE OF HgCdTe PHOTODIODES FOR HETERODYNE APPLICATION	297
Herbert R. Kowitz	
23. EXTENDING THE OPERATING TEMPERATURE, WAVELENGTH AND FREQUENCY RESPONSE OF HgCdTe HETERODYNE DETECTORS	309
D. L. Spears	
24. INFRARED HETERODYNE RECEIVERS WITH IF RESPONSES APPROACHING 5 GHz	327
J. M. Wolczok and B. J. Peyton	
25. FAR INFRARED HETERODYNE SYSTEMS	341
P. E. Tannenwald	
26. BULK SUBMILLIMETER-WAVE MIXERS: STRAIN AND SUPERLATTICES	353
M. M. Litvak and H. M. Pickett	
27. OPTICAL CONSIDERATIONS IN INFRARED HETERODYNE SPECTROMETER DESIGN	365
T. Kostiuik, M. J. Mumma, and D. Zipoy	
28. RF SPECTROMETERS FOR HETERODYNE RECEIVERS	373
D. Buhl and M. J. Mumma	
29. ACOUSTO-OPTIC SPECTROMETER FOR RADIO ASTRONOMY	385
Gordon Chin, David Buhl, and Jose M. Florez	

SESSION V - LOCAL OSCILLATORS B
Chairman: Aram Mooradian

30. MICROWAVE TUNABLE LASER SOURCE: A STABLE, PRECISION TUNABLE HETERODYNE LOCAL OSCILLATOR	399
Glen W. Sachse	

*Papers 20 to 39 are presented under separate cover.

31. SUBMILLIMETER LOCAL OSCILLATORS FOR HETERODYNE SPECTROSCOPY	417
Edward J. Danielewicz, Jr.	
32. SUB MM LASER LOCAL OSCILLATORS: DESIGN CRITERIA AND RESULTS	449
Gerhard A. Koepf	
33. DESIGN CONSIDERATIONS FOR OPTICAL HETERODYNE RECEIVERS:	
A REVIEW	461
John J. Degnan	

SESSION VI - COHERENT APPLICATIONS

Chairman: Frank Goodwin

34. MIXING EFFICIENCY IN HETERODYNE SYSTEMS	487
David Fink	
35. SPATIAL FREQUENCY RESPONSE OF AN OPTICAL HETERODYNE RECEIVER	495
Carl L. Fales and Don M. Robinson	
36. HIGH PRESSURE GAS LASER TECHNOLOGY FOR ATMOSPHERIC REMOTE SENSING	511
A. Javan	
37. AN AIRBORNE DOPPLER LIDAR	529
Charles A. DiMarzio and James W. Bilbro	
38. SURFACE RELIEF STRUCTURES FOR MULTIPLE BEAM LO GENERATION WITH HETERODYNING DETECTOR ARRAYS	541
Wilfrid B. Veldkamp	
39. DIAL WITH HETERODYNE DETECTION INCLUDING SPECKLE NOISE:	
AIRCRAFT/SHUTTLE MEASUREMENTS OF O ₃ , H ₂ O, and NH ₃	
WITH PULSED TUNABLE CO ₂ LASERS	557
Philip Brockman, Robert V. Hess, Leo D. Staton, and Clayton H. Bair	

CORRESPONDENCE

HETERODYNE SIGNAL-TO-NOISE RATIOS IN ACOUSTIC MODE SCATTER- ING EXPERIMENTS	569
William R. Cochran	
PROGRAM COMMITTEE	573
ATTENDEES	575

THE PHYSICS OF HETERODYNE DETECTION IN THE FAR-INFRARED: TRANSITION
FROM ELECTRIC-FIELD TO PHOTON-ABSORPTION DETECTION IN A SIMPLE SYSTEM*

Malvin Carl Teich
Columbia Radiation Laboratory
Department of Electrical Engineering
Columbia University
New York, New York 10027

SUMMARY

After briefly reviewing the history of heterodyne detection from the radiowave to the optical regions of the electromagnetic spectrum, we focus on the submillimeter/far-infrared by investigating the transition from electric-field to photon-absorption detection in a simple system. The response of an isolated two-level detector to a coherent source of incident radiation is calculated for both heterodyne and video detection. When the processes of photon absorption and photon emission cannot be distinguished, the relative detected power at double- and sum-frequencies is found to be multiplied by a coefficient, which is less than or equal to unity, and which depends on the incident photon energy and on the effective temperature of the system.

INTRODUCTION

Heterodyne detection has a long and august history in the annals of electrical engineering, reaching back to the earliest years of the century. The term has its roots in the Greek words "heteros" (other) and "dynamis" (force).

In 1902, Reginald Fessenden was awarded a patent (1) "relating to certain improvements ... in systems where the signal is transmitted by [radio]waves differing in period, and to the generation of beats by the waves and the employment of suitable receiving apparatus responsive only to the combined action of waves corresponding in period to those generated" The subsequent realization that one of these waves could be locally generated (the development of the local oscillator) provided a substantial improvement in system performance. Practical demonstrations of the usefulness of the technique were carried out between the "Fessenden stations" of the U.S. Navy at Arlington (Virginia) and the Scout Cruiser Salem, between the Salem and the Birmingham (1910), and at the National Electric Signaling Company. In

* Work supported by the Joint Services Electronics Program (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract No. DAAG29-79-C-0079.

1913, John Hogan provided a thoroughly enjoyable account of the development, use, and performance of the Fessenden heterodyne signaling system in the first volume of the Proceedings of the IRE (2).

It was not long thereafter, in 1917, that Edwin H. Armstrong of the Department of Electrical Engineering at Columbia University carried out a thorough investigation of the heterodyne phenomenon occurring in the oscillating state of the "regenerative electron relay" (3). A major breakthrough in the field, the development of the superheterodyne receiver, was achieved by Armstrong in 1921 (4), and this famous invention is now used in systems as diverse as household AM radio receivers and microwave Doppler radars. The prefix "super" refers to "the super-audible frequency that could be readily amplified".

In the succeeding years, the application of heterodyne and superheterodyne principles followed the incessant march toward higher frequencies that culminated in the remarkable developments in microwave electronics about the time of World War II.

The marriage of heterodyning and the optical region took place in 1955. In a now classic experiment, Forrester, Gudmundsen, and Johnson (5) observed the mixing of two Zeeman components of a visible (incoherent) spectral line in a specially constructed photomultiplier tube. With the development of the laser, optical heterodyning became considerably easier to observe and was studied in 1962 by Javan, Ballik, and Bond (6) at $1.15 \mu\text{m}$ using a He-Ne laser, and by McMurtry and Siegman (7) at 6943 \AA using a ruby laser.

The development of new transmitting and receiving components in the middle infrared region of the electromagnetic spectrum led Teich, Keyes, and Kingston (8) in 1966 to perform a heterodyne experiment using a CO_2 laser at $10.6 \mu\text{m}$ in conjunction with a copper-doped germanium photoconductive detector operated at 4°K . Subsequent experiments with lead-tin selenide photovoltaic detectors confirmed the optimal nature of the detection process (9)-(11). The state of the art in applying these results to coherent infrared radar was reviewed by Kingston in 1977 (12). Elbaum and Teich (13) have recently shown the importance of properly choosing the performance measure for a heterodyne system. This can sometimes be critical to avoid faulty estimates of signal and noise levels leading to inconsistent or incorrect results, as pointed out by a number of authors (8), (14).

The foregoing discussion has dealt with more-or-less idealized systems; it must be kept in mind that there are a variety of effects that can alter heterodyne performance in important ways. Two examples were readily provided at this conference: Charles Townes discussed atmospheric perturbations of phase coherence in large diameter heterodyne receivers, and M. J. Mumma and T. Kostiuik discussed the effects of various sources of noise on system performance.

Finally, we point out that a number of system configurations employing different forms of nonlinear heterodyne detection have been proposed for various applications (15), (16). These make use of multiple frequencies,

nonlinear detectors, and/or correlation schemes. All are aimed at increasing the signal-to-noise ratio in situations for which usual operating conditions are relaxed in particular ways. One of these schemes (17), (15), using a two-frequency transmitter, is currently under consideration for possible application in over-water pilot rescue missions where the use of a light-chopping mechanism is interdicted because of its interruptive nature (private communication with W. Tanaka, Naval Weapons Center).

In the following section, we relax the assumption that heterodyne detection takes place by means of photon absorption, as it does in the middle infrared and optical regions (18), (19), or by means of electric-field detection, as it does in the radiowave and microwave regions. This will enable us to examine the heterodyne detection process in the submillimeter/far-infrared region of the spectrum in a simple way. The reader is cautioned that the treatment is heuristic in nature, and is only intended to provide a qualitative description of the underlying detection process.

TRANSITION FROM ELECTRIC-FIELD TO PHOTON-ABSORPTION DETECTION IN A SIMPLE SYSTEM

The increasing importance of heterodyning and electric-field detection in the submillimeter, infrared, and optical (20), (21) has encouraged us to examine the operation of such systems in relation to the more conventional photon-absorption detector (22)-(24).

For heterodyne mixing with coherent signals in a two-level detector, a simple argument indicates that double- and sum-frequency terms in the detected power are multiplied by the factor $\text{sech}(h\nu/2kT_e)$, which varies smoothly from unity in the electric-field detection regime to zero in the photon-absorption detection regime. This factor depends on both the incident photon energy $h\nu$ and on the effective excitation energy of the (two-level) detector, kT_e . It is observed that these terms only appear when it cannot be determined whether photon absorption or photon emission has taken place. Difference-frequency signals, on the other hand, arise from our inability to determine from which beam a photon is absorbed in a heterodyne experiment (18), (19).

We consider a simple hypothetical two-level system with an effective temperature T_e . It is assumed that the system responds linearly to the incident radiation intensity, and that its interaction with the field is sufficiently weak such that the state of the field is not perturbed by the presence of the detector.

We label the initial and final states of the system as $|\alpha\rangle$ and $|\Omega\rangle$ and of the radiation field as $|i\rangle$ and $|f\rangle$, respectively. For an electric-dipole transition, the transition probability W_{fi} (which is related to the detected power) is given approximately by

$$W_{fi} = |\langle f | \Omega | e q (E^+ + E^-) | \alpha i \rangle|^2, \quad (1)$$

where e is the electronic charge, q is the detector coordinate, and E^+ and E^- are the positive- and negative-portions of the electric-field operator, respectively. Since E^+ corresponds to photon absorption or annihilation, and E^- corresponds to photon emission or creation, the transition probability may be written as

$$W_{fi} = |\langle fu | eq a_\ell E^+ | \ell i \rangle + \langle f \ell | eq a_u E^- | u i \rangle|^2, \quad (2)$$

where $|\ell\rangle$ and $|u\rangle$ represent the lower and upper states of the two-level system, respectively. This equation assumes that the atomic system is, in general, in a superposition state. Using microscopic reversibility, the quantity $|\langle u | eq | \ell \rangle|^2$, which represents the quantum efficiency κ , may be factored out of Eq. (2). We then sum over the final states of the field (23), which are not observed, to obtain

$$\begin{aligned} W \propto & \langle i | a_\ell^* a_\ell E^- E^+ | i \rangle + \langle i | a_u^* a_u E^+ E^- | i \rangle \\ & + \langle i | a_\ell^* a_u E^- E^- + a_\ell a_u^* E^+ E^+ | i \rangle. \end{aligned} \quad (3)$$

The normally ordered first term corresponds to stimulated absorption, the antinormally ordered second term corresponds to photon emission (25), and the third is an interference term.

We now assume that before the interaction, the probability amplitudes of the two possible states, a_ℓ and a_u , were related by the Boltzmann factor, with lower and upper level energies represented by E_ℓ and E_u , respectively, and with excitation energy kT_e defining the effective temperature of the detector. Thus,

$$|a_u|^2 / |a_\ell|^2 = \exp[-(E_u - E_\ell) / kT_e], \quad (4)$$

yielding

$$a_\ell = (1 + e^{-x})^{-1/2} e^{i\phi} \quad (5)$$

and

$$a_u = e^{-x/2} (1 + e^{-x})^{-1/2} e^{i\theta}, \quad (6)$$

with $x \equiv h\nu / kT_e$ and $e^{i\theta}$, $e^{i\phi}$ representing phase factors. Then, generalizing to an arbitrary radiation field represented by the density operator ρ , we cavalierly obtain the expression

$$W \propto \left\{ [e^{x/2} \text{sech}(x/2)] \text{tr}\{\rho E^- E^+\} + [e^{-x/2} \text{sech}(x/2)] \text{tr}\{\rho E^+ E^-\} \right. \\ \left. + [\text{sech}(x/2)] \text{tr}\{\rho (e^{-i\gamma} E^- E^- + e^{i\gamma} E^+ E^+) \} \right\}, \quad (7)$$

with $\gamma \equiv \phi - \theta$.

Considering ideal heterodyne detection, i.e., two parallel, monochromatic, and coherent waves of frequencies ν_1 and ν_2 impinging normally on the detector, and neglecting spontaneous emission so that the antinormally ordered and the normally ordered terms are equal in magnitude (25), the first two terms above generate dc and difference-frequency signals, while the third term contributes double- and sum-frequency signals. This may be clearly seen by explicitly rewriting Eq. (7) as

$$W \propto |\epsilon_1^0|^2 + |\epsilon_2^0|^2 + 2|\epsilon_1^0||\epsilon_2^0| \cos[2\pi(\nu_1 - \nu_2)t + (\beta - \alpha)] \\ + [\text{sech}(h\nu/2kT_e)] \left\{ |\epsilon_1^0|^2 \cos(4\pi\nu_1 t - 2\alpha - \gamma) + |\epsilon_2^0|^2 \cos(4\pi\nu_2 t - 2\beta - \gamma) \right. \\ \left. + 2|\epsilon_1^0||\epsilon_2^0| \cos[2\pi(\nu_1 + \nu_2)t - (\alpha + \beta) - \gamma] \right\} \quad (8)$$

where $\epsilon_1^0 = |\epsilon_1^0| e^{i\alpha}$ represents the complex electric-field amplitude of the constituent field with frequency ν_1 and phase α . Double- and sum-frequency terms in the heterodyne signal are therefore multiplied by the factor $\text{sech}(h\nu/2kT_e)$.

For $h\nu/kT_e \rightarrow 0$, this factor approaches 1 and the classical electric-field heterodyne signal obtains

$$W_{\text{elec}} \propto \left[|\epsilon_1^0| \cos(2\pi\nu_1 t - \alpha - \frac{\gamma}{2}) + |\epsilon_2^0| \cos(2\pi\nu_2 t - \beta - \frac{\gamma}{2}) \right]^2. \quad (9)$$

For $h\nu/kT_e \rightarrow \infty$, $\text{sech}(h\nu/2kT_e) \rightarrow 0$ and the photon-absorption (optical) heterodyne signal obtains (18), (19),

$$W_{\text{abs}} \propto \left\{ |\epsilon_1^0|^2 + |\epsilon_2^0|^2 + 2|\epsilon_1^0||\epsilon_2^0| \cos[2\pi(\nu_1 - \nu_2)t + (\beta - \alpha)] \right\}. \quad (10)$$

A graphical presentation of the function $\text{sech}(h\nu/2kT_e)$ vs. $(h\nu/kT_e)$ is provided in Fig. 1.

The result for the heterodyne case is easily reduced to the direct detection (video) case for a coherent signal by setting $|\epsilon_2^0| = 0$, which yields

$$W_{\text{dir}} \propto |\epsilon_1^0|^2 \left\{ 1 + [\text{sech}(h\nu/2kT_e)] [\cos(4\pi\nu_1 t - 2\alpha - \gamma)] \right\}. \quad (11)$$

Thus, double-frequency intensity fluctuations are discerned for electric-field direct detectors, while they are suppressed for photon-absorption direct detectors which respond simply as $|\epsilon_1^0|^2$.

CONCLUSION

The photon-absorption detector is, by definition, initially in its ground state and functions by the annihilation of a single (in general nonmonochromatic) photon (18). The presence of the difference-frequency signal is understood to arise from our inability to determine from which of the two constituent beams the single photon is absorbed (19). The two-level electric-field detector, on the other hand, has equal probability of being in the lower and in the upper state, so that the pure processes of photon annihilation and photon absorption occur with equal likelihood, and we must add the effects of both. When we are unable to determine which of these processes is occurring, we must add amplitudes rather than squares of amplitudes, thereby allowing interference to occur. Any attempt, in this case, to determine whether photon emission or photon absorption takes place would randomize the phase γ , and thereby wash out the sum- and double-frequency components. In general, then, a photon incident on a video detector induces upward and downward transitions with different probabilities. This, in turn, creates a quantum-mechanical probability density (and charge distribution) that varies in time, producing a current at the incident radiation frequency. The power absorbed then contains a double-frequency signal. For heterodyne detection, in the general case, sum-frequency signals are observed as well.

The foregoing heuristic model yields a simple result for an idealized two-level system. Replacing the operator eq by the non-relativistic Hamiltonian $(\hat{p} - e\hat{A})^2/2m$, where \hat{p} and \hat{A} represent the momentum and vector-potential operators, respectively, would allow transitions more general than electric-dipole, and absorptions of more than one quantum, to occur. A rigorous treatment should consider a collection of such systems (as a model for a bulk photodetector or metal antenna), in the presence of a surrounding reservoir, and should be carried out using the density matrix formalism. It is expected that the effects described here are important in a broad range of systems, including the Josephson detector (26), (27).

Tucker (28) has recently carried out a rigorous analysis of quantum-limited detection in tunnel junction mixers. He demonstrates that nonlinear tunneling devices are predicted to undergo a transition from energy detectors to photon counters at frequencies where the photon energy becomes comparable to the voltage scale of the dc nonlinearity. It is apparent that the character of this result is closely related to the discussion presented here.

REFERENCES

- (1) Fessenden, R. A.: Wireless Signaling. U. S. Patent No. 706,740, August 12, 1902.
- (2) Hogan, J.L.: The Heterodyne Receiving System, and Notes on the Recent Arlington-Salem Tests. Proc. IRE, vol. 1, pp. 75-102, July 1913.
- (3) Armstrong, E. H. : A Study of Heterodyne Amplification by the Electron Relay. Proc. IRE, vol. 5, pp. 145-168, April 1917.
- (4) Armstrong, E. H. : A New System of Short Wave Amplification. Proc. IRE, vol. 9, pp. 3-27, February 1921.
- (5) Forrester, A. T.; Gudmundsen, R. A.; and Johnson, P.O.: Photoelectric Mixing of Incoherent Light. Phys. Rev., vol. 99, pp. 1691-1700, September 1955.
- (6) Javan, A; Ballik, E. A.; and Bond, W. L.: Frequency Characteristics of a Continuous-Wave He-Ne Optical Maser. J. Opt. Soc. Amer., vol. 52, pp. 96-98, January 1962.
- (7) McMurtry, B. J.; and Siegman, A. E. : Photomixing Experiments with Ruby Optical Maser and Traveling Wave Microwave Phototubes. Appl. Opt., vol. 1, pp. 51-53, January 1962.
- (8) Teich, M. C.; Keyes, R. J.; and Kingston, R. H.: Optimum Heterodyne Detection at 10.6 μm in Photoconductive Ge:Cu. Appl. Phys. Letters, vol. 9, pp. 357-360, 15 November 1966.
- (9) Teich, M. C. : Infrared Heterodyne Detection. Proc. IEEE, vol. 56, pp. 37-46, January 1968 [reprinted in Infrared Detectors, Hudson, R. D., Jr., and Hudson, J. W. Eds. (Dowden, Hutchinson, and Ross, Stroudsburg, 1975)].
- (10) Teich, M. C. : Homodyne Detection of Infrared Radiation from a Moving Diffuse Target. Proc. IEEE, vol. 57, pp. 786-792, May 1969.
- (11) Teich, M. C. : Coherent Detection in the Infrared. In Semiconductors and Semimetals, Willardson, R. K., and Beer, A. C. Eds. (Academic, New York, 1970), vol. 5, Infrared Detectors, ch. 9, pp. 361-407.
- (12) Kingston, R. H.: Coherent Optical Radar. Optics News, vol. 3, no. 3, pp. 27-31, Summer 1977.
- (13) Elbaum, M; and Teich, M. C.: Heterodyne Detection of Random Gaussian Signals in the Optical and Infrared: Optimization of Pulse Duration. Opt. Commun., vol. 27, pp. 257-261, November 1978.

- (14) Brockman, P.; Hess, R. V.; Staton, L. D.; and Bair, C. H.: DIAL With Heterodyne Detection Including Speckle Noise: Aircraft/Shuttle Measurements of O_3 , H_2O , NH_3 With Pulsed Tunable CO_2 Lasers. Heterodyne Systems and Technology, NASA CP-2138, 1980. (Paper 39 of this compilation.)
- (15) Teich, M. C.: Nonlinear Heterodyne Detection. In Topics in Applied Physics, Keyes, R. J. Ed. (Springer-Verlag, Berlin/Heidelberg, 1977), vol. 19, Optical and Infrared Detectors, ch. 7, pp. 229-300.
- (16) Teich, M. C.: Heterodyne Correlation Radiometry. Opt. Engineering, vol. 17, pp. 170-175, March 1978.
- (17) Teich, M. C. : Three-Frequency Heterodyne System for Acquisition and Tracking of Radar and Communications Signals. Appl. Phys. Letters, vol. 15, pp. 420-423, December 1969.
- (18) Teich, M. C. : Field-Theoretical Treatment of Photomixing. Appl. Phys. Letters, vol. 14, pp. 201-203, March 1969.
- (19) Teich, M. C. : Quantum Theory of Heterodyne Detection. In Proc. Third Photoconductivity Conf., Pell, E. M. Ed. (Pergamon, New York, 1971), pp. 1-5.
- (20) Hocker, L. O.; Sokoloff, D. R.; Daneu, V.; Szűke, A.; and Javan, A.: Frequency Mixing in the Infrared and Far-Infrared Using a Metal-to-Metal Point Contact Diode. Appl. Phys. Letters, vol. 12, pp. 401-402, June 1968.
- (21) Abrams, R. L.; and Gandrud, W. B.: Heterodyne Detection of $10.6\text{-}\mu$ Radiation by Metal-to-Metal Point Contact Diodes. Appl. Phys. Letters, vol. 17, pp. 150-152, August 1970.
- (22) Mandel, L.: Fluctuations of Photon Beams: The Distribution of the Photo-Electrons. Proc. Phys. Soc. (London), vol. 74, pp. 233-243, 1959.
- (23) Glauber, R. J.: The Quantum Theory of Optical Coherence. Phys. Rev., vol. 130, pp. 2529-2539, June 1963; Coherent and Incoherent States of the Radiation Field. Phys. Rev., vol. 131, pp. 2766-2788, September 1963.
- (24) Kelley, P. L.; and Kleiner, W. H.: Theory of Electromagnetic Field Measurement and Photoelectron Counting. Phys. Rev., vol. 136, pp. A316-A334, October 1964.
- (25) Mandel, L: Antinormally Ordered Correlations and Quantum Counters. Phys. Rev., vol. 152, pp. 438-451, December 1966.
- (26) McDonald, D. G.; Risley, A. S.; Cupp, J. D.; Evenson, K. M.; and Ashley, J. R.: Four-Hundredth-Order Harmonic Mixing of Microwave and Infrared Laser Radiation Using a Josephson Junction and a Maser. Appl. Phys. Letters, vol. 20, pp. 296-299, April 1972.

- (27) Zimmerman, J. E.: Heterodyne Detection with Superconducting Point Contacts and Enhanced Heterodyne Signals from Tightly Coupled Contacts. J. Appl. Phys., vol. 41, pp. 1589-1593, March 1970.
- (28) Tucker, J. R.: Quantum Limited Detection in Tunnel Junction Mixers. IEEE J. Quantum Electron., vol. QE-15, pp. 1234-1258, November 1979.

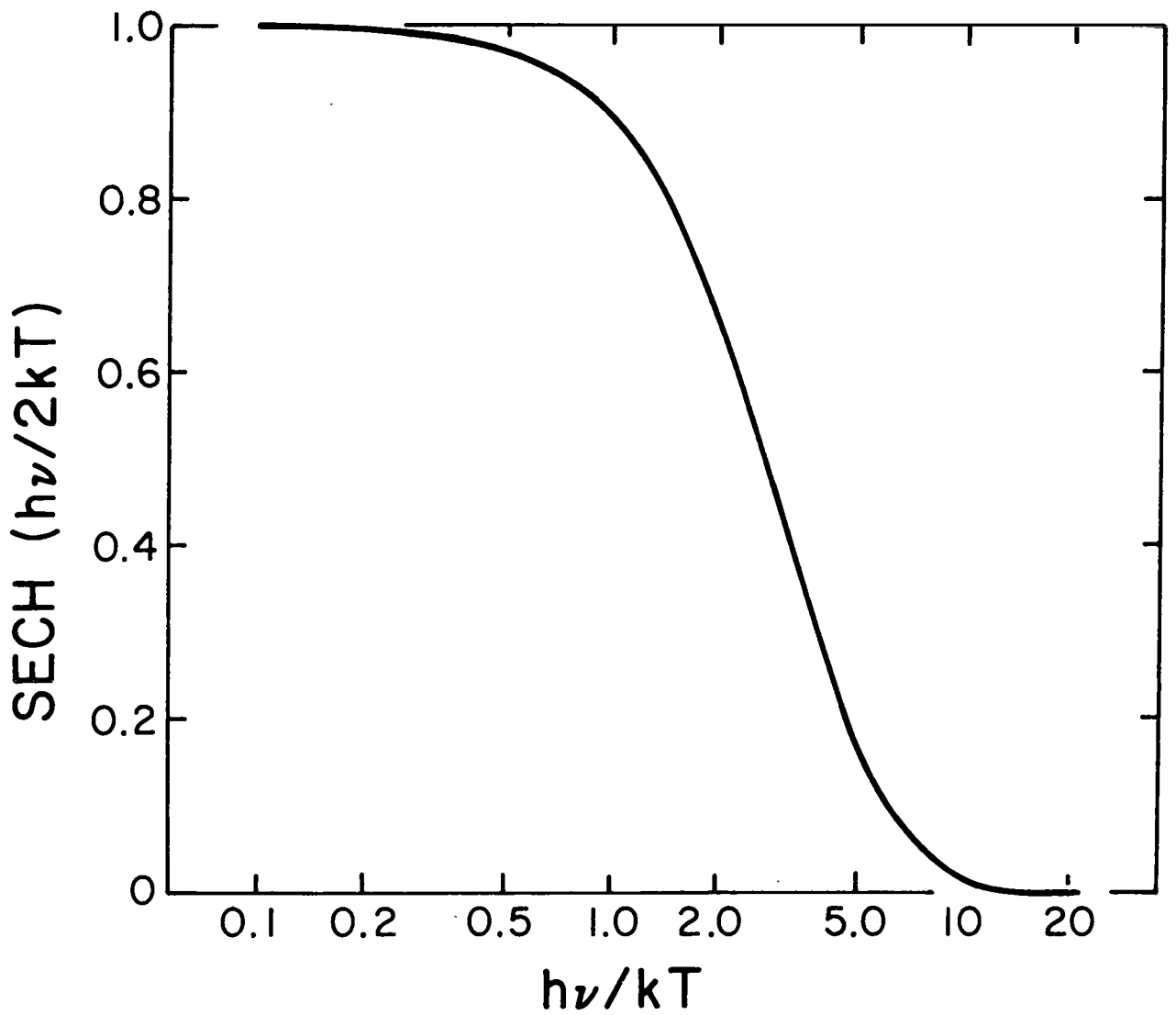


Figure 1.- Factor $\text{sech}(h\nu/2kT)$ vs. $h\nu/kT$. This quantity appears as a coefficient in the absorption/emission interference term.

INFRARED HETERODYNE SPECTROSCOPY IN ASTRONOMY*

Albert Betz
Department of Physics
University of California, Berkeley

SUMMARY

A heterodyne spectrometer has been constructed and applied to problems in infrared astronomical spectroscopy. The instrument offers distinct observational advantages for the detection and analysis of individual spectral lines at Doppler-limited resolution. Observations of carbon dioxide in planetary atmospheres and ammonia in circumstellar environments demonstrate the substantial role that infrared heterodyne techniques will play in the astronomical spectroscopy of the future.

INTRODUCTION

For astronomical observations which require Doppler-limited resolution and high absolute frequency accuracy, laser heterodyne spectroscopy offers to the infrared the technical advantages long enjoyed at radio frequencies. The vibrational-rotational spectra of planetary and circumstellar molecules can now be studied at the same high resolution commonly used for the pure rotational and other low-energy transitions falling in the microwave domain. In the infrared, the Doppler-limited resolution of heterodyne spectroscopy has made possible new types of observations. Examples are: studies of exact lineshapes of CO_2 absorption profiles in the atmospheres of Mars and Venus to derive the vertical pressure-temperature structure of the atmosphere, accurate measurements of wind velocities in the stratosphere and mesosphere of Venus from the Doppler shifts of narrow CO_2 lines, and investigations of the gas dynamics of circumstellar molecules by the detection of the $10\text{ }\mu\text{m}$ absorption lines of NH_3 .

In heterodyne spectroscopy, whether radio or infrared, the source radiation collected through a telescope is mixed with a stable local oscillator signal in a nonlinear element (usually a device with a square-law response to the electric field), and the resulting difference frequencies are analyzed in a bank of radio frequency (RF) filters and power detectors. In a properly designed system, the frequency calibration and stability of the local oscillator (such as a klystron in the radio or a laser in the infrared) determines the absolute frequency calibration and ultimate frequency resolution capability of the entire spectrometer. In practice, the frequency widths of the individual RF filter channels are usually chosen to resolve adequately a Doppler-limited profile. The total number of available channels, and hence the total spectral range visible at one time, may be dictated either by economics or the intermediate

*This work supported in part by NASA Grant NGR 05-003-452.

frequency (IF) bandwidth available from the mixer. Doppler-limited spectroscopy usually requires resolving powers better than 10^6 , and so whereas channel widths of 300 kHz may be adequate in a 1 mm wavelength receiver, channels as wide as 30 MHz would be more practical for a 10 μ m spectrometer. The following sections discuss the implementation and use of such a heterodyne spectrometer for the 10 μ m band, with initial emphasis on the local oscillator and mixer characteristics of the instrument.

INSTRUMENT

Figure 1 shows a simplified schematic of the spectrometer. The converging beam from the telescope enters on the left, comes to a focus, and then passes through a coated NaCl beamsplitter. (The telescope used for all these observations is either the 1.5 or 0.75 meter apertures of the McMath Solar Telescope at Kitt Peak National Observatory in Tucson.) The vertically-polarized output from a CO₂ laser is first matched to the divergence of the signal beam and then partially reflected by the beamsplitter. By expanding the Gaussian laser beam before the beamsplitter and then selecting only the center, a nearly uniform intensity distribution can be obtained to match with the nearly uniform transverse intensity distribution of the telescope beam. In practice, the "lens" at the output of the laser is a pair of small mirrors which perform the equivalent beam-matching function illustrated in the diagram. The second germanium lens just before the detector is as shown, however, and is used at approximately f/3 in order to achieve a small spot size (~ 60 μ m) at an optimum point on the photodetector (~ 150 μ m diameter). Only about 0.5 to 0.8 mW of laser power is actually incident on the detector, an HgCdTe photodiode developed by David Spears of Lincoln Laboratory. With a low noise preamplifier following the mixer, useful bandwidths to 2 GHz can easily be obtained with these photodetectors, although the particular ones we have are normally rated to ~ 1.5 GHz. The question of defining IF bandwidth in a real application will be discussed later. The amplified IF output is then directed into a second RF mixer which converts a selected 1280 MHz wide segment into a 64x20 MHz filter bank. The undesired image response of the 2nd mixer is rejected by filtering before the mixer. The direct RF to IF isolation of this mixer is measured to exceed 26 dB over the output band. (Future designs of this subsystem will employ a more complicated but versatile combination of 2nd and 3rd local oscillators.) The 2nd LO is also used to track out changes of source Doppler velocity (such as rotation of the Earth ~ 3 MHz/hr) during the course of observations. For planetary observations a different filter bank of 40x5 MHz filters (200 MHz total) is used to provide finer resolution. The detected outputs from the filter-bank channels are simultaneously integrated in a following set of analog integrators and the net outputs periodically sampled by a computer. On weak sources, repeated integrations of 8 min interspaced by 2 min calibrations on a blackbody source may accumulate for several hours. Not shown in this figure are sky-chopper foreoptics at the telescope focus which alternate the field-of-view of the detector on and off the source at 150 Hz. The chopped signals from the filter bank are synchronously demodulated at this frequency just before integration.

FRONT END COMPONENTS

Laser

All our observational work to date has been accomplished with fixed-frequency CO₂ lasers, which are step-tunable to discrete vibrational-rotational transitions of various CO₂ isotopes over the 9 to 12 μ m band. The particular model now in use is based on the Lincoln Laboratory design of Charles Freed in using a semi-sealed discharge tube, invar stabilizing rods, and grating tuning control with zero-order output coupling.¹ This design maximizes the Q of the 1½ m laser cavity and permits laser oscillation on J-values up to ~60 on the stronger isotopes of CO₂. Output powers up to several watts are still obtained on the stronger lines.² Aside from continuous tunability, this laser offers all the features desired in a local oscillator: more than adequate power in a clean single mode, freedom from AM and FM perturbations, reasonable physical size and efficiency, and accurate absolute frequency calibration for all discrete oscillation frequencies.² In fact, the lack of continuous tunability and hence uncertainty in the oscillation frequency may be viewed as a positive feature for many spectroscopic applications, in that the laser frequency is auto-calibrated either to better than 1 part in 10⁷ accuracy with the discharge impedance sensing technique,³ or to better than 1 part in 10⁹ absolute accuracy with the saturated resonance fluorescence technique.⁴ Consequently, a separate involved calibration mechanism is not necessary, as would be the case with tunable diode lasers, for example. The completeness of spectral coverage in the 9 to 12 μ m region can be estimated by counting the total number of laser lines available from isotopic CO₂ and ¹⁴N₂¹⁶O lasers in the sequence and "hot" bands as well as the more conventional laser bands. The average spacing between lines is ~3 GHz, which is also close to the IF output bandwidth available from state-of-the-art infrared photomixers. This is not to say that a laser line is always available to convert a target frequency into the center of the available IF band, but that, more often than not, a laser line can be found in the 9 to 12 μ m range so that practical observations may be attempted. An added influencing factor aiding the use of a fixed-frequency LO is that the Doppler frequency of the target line will vary with the orbital motion of the Earth, and shifts as large as ±3 GHz (±30 km/s) can be expected at 10 μ m, depending on the direction to the source. Even for the least favorably placed stellar sources, during at least one month of the year a specified line will usually be observable with a CO₂ or N₂O laser. The more favorable planetary sources, the planets Mars and Venus with rich CO₂ atmospheres, are trivially observable from a Doppler viewpoint, as the Doppler shifts from motion with respect to the Earth do not exceed 1.5 GHz (15 km/s). In this latter application, where CO₂ itself is the target molecule, the CO₂ laser is of course ideally suited as the local oscillator. In those other cases where the target line is not a laser transition, such as the NH₃ lines seen in stars, then the offset frequency between the target line and the laser must be determined by laboratory spectroscopy. Aside from the heterodyne measurements done in our own laboratory on specific lines of NH₃, OCS, C₂H₄, and SiH₄, both heterodyne and a number of other laser-related techniques have been applied to the 10 μ m spectroscopy of astrophysically important molecules with ~3 MHz accuracy.⁵⁻¹¹

Detector/Preamplifier

The photomixer is a reversed-biased HgCdTe photodiode with an $\sim 2 \times 10^{-4} \text{ cm}^{-2}$ active area and an effective AC quantum efficiency of $\sim 35\%$ at $10.6 \text{ } \mu\text{m}$.¹² The LO power level required for quantum-noise (shot-noise) limited operation depends, of course, on the noise temperature of the first IF amplifier stage. With conventional bipolar amplifiers featuring an $\sim 250 \text{ K}$ noise temperature over the 100 to 1500 MHz band, laser power levels of about 0.5 to 0.8 mW are required on the detector to raise the laser-induced shot noise from the photomixer to several times the amplifier noise level. LO power levels much above 1 mW on the detector risk inducing saturation and a concomitant decrease in effective quantum efficiency. An operational definition of the detector bandwidth for quantum-noise-limited performance can be stated as the IF frequency where the LO induced shot noise and amplifier noise contributions are equal. Operational bandwidths exceeding the detector's -3 dB bandwidth are thus possible. For the Lincoln Laboratory photomixers that we use, the output bandwidth is $\sim 1500 \text{ MHz}$ for operation with a bipolar preamplifier with a 250 K noise temperature. Obviously, a reduction in amplifier noise at high intermediate frequencies can extend the usable bandwidth of the photodetector. Any drop in IF system gain with detector rolloff can be compensated in later stages of amplification where the noise temperature factor is not critical. The current availability of low-noise GaAs field effect transistors permits an optimized preamplifier to be constructed to match the photomixer response characteristics. An amplifier based on designs of the Radio Astronomy Laboratory at Berkeley is now in use on the spectrometer. It features a reactively-tuned 2-stage design with about 20 dB gain and good input/output match ($\text{VSWR} \leq 2.0$) from 500 to 2000 MHz. The noise temperature minimum is adjusted to fall at $\sim 1600 \text{ MHz}$ so as to provide the best performance at a frequency where the photomixer response is already falling off. When operated at room temperature, the amplifier has a noise temperature of 60 K at 1600 MHz, 80 K at 2000 MHz, and 90 K at 1000 MHz. Below 1 GHz, the noise temperature rises rapidly to $\sim 250 \text{ K}$ at 500 MHz. When the preamplifier is cooled to 77 K, as depicted in Figure 1, the noise temperatures fall to about half of their respective room temperature values. Even for a warm amplifier, however, the detector/preamp system operates in the quantum noise limit to beyond 2 GHz, the limit of our measurement system. Although this relatively slight increase of 30% in operating bandwidth may appear small and not worth the additional effort required, in practice it significantly improves the spectrometer's usefulness, in that wide spectral lines can be centered at a higher intermediate frequency and thus avoid the lineshape interpretation problems associated with side-band foldover about the local oscillator frequency. In addition, several lines of circumstellar NH_3 have been detected with the 2 GHz system which are unreachable with only a 1.5 GHz IF. For astronomical observations, more benefits will come from increased output bandwidths in photomixers than from improvements in quantum efficiency, although, of course, both are desired. But, even as improved detectors with 3 GHz bandwidths (-3 dB) are fabricated in the near future, cooled FET preamps will still be desired to further extend operational bandwidths. Currently, a cooled preamp with an 0.1 to 4.0 GHz response is under development at Berkeley. This amplifier will be integrated with the photodiode package to eliminate intervening cabling and the attendant problems associated with VSWR variations.

Another obvious advantage of a low noise preamp is the reduction in LO drive level required to raise the laser-induced shot noise above the preamplifier noise contribution. Cooled FET preamps should allow receivers using tunable diode lasers of modest single mode power ($\sim 100 \mu\text{W}$) to achieve quantum-noise-limited performance over IF bandwidths approaching the detector's -3 dB response.

OPERATIONAL CHARACTERISTICS

The sensitivity of the spectrometer on the telescope may be computed directly from the throughputs of the various components. At a wavelength of $10.5 \mu\text{m}$, 92% of the incident signal flux is transmitted through the terrestrial atmosphere, 77% through the 5 mirrors of the solar telescope, 50% average through the sky-chopper, and 75% through the table optics. Of the remaining flux, 35% produces an effective IF signal current in the photomixer. Furthermore, only signal radiation in the polarization of the laser is detected, which may be regarded as an additional 50% loss for unpolarized sources. The net system detection efficiency of only 5%, however, is still adequate for useful astronomical work, as will be seen in the next section.

Spectra are generated by integrating for 4 minutes with the source in the positive beam of the sky-chopper and then for 4 minutes with the source in the negative beam. The difference between these integrations is the total signal with integrator offsets cancelling. Laboratory blackbodies at ambient and an elevated temperature (228°C) are then interposed in the two beams for the calibration cycle and the process repeated, except that each integration is for 1 minute. The final spectra is produced by normalizing the signal by the blackbody integration. Previous calibrations of telescope and atmospheric transmission allow an absolute signal level to be inferred. It is important that the laser power on the photodetector remain constant until the calibration cycle is completed. Otherwise, changes in the photodiode impedance will not allow gain variations from standing waves between mixer and preamp to be completely normalized by the calibration cycle, and a distorted spectra may result.

The infrared transmission of the terrestrial atmosphere can be determined from spectra of the Moon or the planet Mercury. The Sun may also be used as a blackbody for spectral calibration of the terrestrial atmosphere if accuracies no better than $\sim 1\%$ are required. Our repeated observations of the intrinsic solar spectrum show numerous unidentified absorption lines throughout the $10 \mu\text{m}$ spectrum. The lines are on the order of 1% deep and 300 to 400 MHz wide, and vary in strength and detail from place to place on the Sun. Possible contributors include the fundamental vibrational-rotational bands of FeO and other metallic oxides; however, the current state of laboratory spectroscopy on these molecules is relatively so poor that positive identifications may not readily be achieved.

ASTRONOMICAL APPLICATIONS

Planets

The first use of infrared heterodyne for astronomical spectroscopy was to measure the lineshapes of individual $^{13}\text{C}^{16}\text{O}_2$ lines in the atmosphere of Mars.¹³ Figure 2 shows a subsequent observation with an improved version of that early spectrometer. The P(16) line of $^{12}\text{C}^{16}\text{O}_2$ in the 10 μm laser band is seen in absorption in the Martian atmosphere. Each channel is 5 MHz wide and overlapping segments of 40 channels were combined to reconstruct the wing of half the line. The equivalent integration time for the spectrum is 40 minutes. The intent of this and similar observations of other CO_2 lines was to model the vertical pressure-temperature structure of the atmosphere of Mars.¹⁴ The data were fitted with a simple 4 parameter model, and the solid curves show the effects of different surface pressure estimates to the fit of the data. The ability of heterodyne spectroscopy to resolve the lineshape permits measurements of atmospheric surface pressure to better than 0.5 millibar (~ 0.4 torr) accuracy. Two additional unexpected features (at the time) are also seen in this profile. The small arrow on the right points to the absorption line of the P(23) transition of $^{12}\text{C}^{16}\text{O}^{18}\text{O}$ in the 10 μm laser band. On the left, the central core of the $^{12}\text{C}^{16}\text{O}_2$ absorption line is seen to be in emission. The width of this Gaussian re-emission component is only ~ 35 MHz (FWHM), which implies a gas kinetic temperature of only ~ 170 K at the altitude of line formation. The intensity of the emission, however, is much stronger than that expected from CO_2 in local thermodynamic equilibrium at this temperature and suggests a nonthermal excitation mechanism for the phenomenon. The detection of similar and stronger re-emission cores in $^{12}\text{C}^{16}\text{O}_2$ absorption lines on Venus strengthens the interpretation that the nonthermal excitation of CO_2 comes from direct solar pumping of short wavelength CO_2 bands.^{15, 16} The emitting CO_2 lies high in the mesospheres of the planets (75 km on Mars, 120 km on Venus) so that collisional de-excitation cannot quench the subsequent 10 μm "fluorescence". Figure 3 illustrates the appearance of the re-emission core of the P(16) line for three different positions on Venus. Sunlight illuminates the planet from the left in this diagram and is strongest at position 1, close to the local noon on the planet. The strength of the 10 μm emission line is clearly seen to be proportional to the incident solar intensity. No emission line is seen from the dark half of the planet at the right; only the relatively flat residual continuum radiation at the center of the very broad $^{12}\text{C}^{16}\text{O}_2$ absorption line is detected. On Venus, the column density of $^{12}\text{C}^{16}\text{O}_2$ is so high that the absorption lines of $^{12}\text{C}^{16}\text{O}_2$ are "saturation-broadened" to be much wider than our photomixer bandwidth. The complete profiles of weaker lines of $^{13}\text{C}^{16}\text{O}_2$ can be readily seen in absorption, however. For Venus, the

Doppler-shifts of both the mesospheric $^{12}\text{C}^{16}\text{O}_2$ emission components (120 km) and the stratospheric $^{13}\text{C}^{16}\text{O}_2$ absorption lines (80 km altitude) have been monitored for several years in order to determine the average wind circulation patterns at these altitudes. The signal-to-noise ratio and absolute laser frequency calibration of the heterodyne spectra permit emission and absorption line shifts to be measured to ~ 2 and 15 m/sec accuracy, respectively. The results indicate an average retrograde circulation of 90 m/sec at stratospheric altitudes and a symmetric subsolar to anti-solar flow as fast as 130 m/sec in the mesosphere.¹⁷ (A complete report on four years of observations will be submitted for publication following planned observations in late Spring, 1980.)

Stars

In 1978, the first application of the spectrometer to stellar spectroscopy was the detection of several absorption lines in the $10\ \mu\text{m}\ \nu_2$ band of ammonia in the gas cloud surrounding a supergiant star called IRC + 10216.¹⁸ Since then, circumstellar ammonia has been detected around a number of supergiant stars of various types, some of which also emit peculiar microwave maser emission from circumstellar OH, H_2O , and SiO molecules.^{19,20} Ammonia is relatively abundant in these sources and is an excellent indicator of the gas dynamics throughout the circumstellar region. The absolute frequency accuracy and resolution of the infrared heterodyne spectrometer permits direct comparisons with similarly accurate microwave data available from radio astronomy on these maser stars. For the non-maser star IRC + 10216, Figure 4 shows an absorption line of $^{14}\text{NH}_3$ in the ground (0,0) rotational state, a level well-populated in the colder ($< 200\ \text{K}$) regions of the outer circumstellar envelope. The integration time on this line was 80 minutes. Laboratory spectroscopy on this transition placed it only 753 MHz below the P(20) laser line of $^{12}\text{C}^{18}\text{O}_2$ in the $10\ \mu\text{m}$ band. Because of the abundance of ground-state ammonia, this particular laser line has turned out to be the "workhorse" LO frequency for the initial detection of NH_3 in stellar sources. The Doppler velocity of the circumstellar gas with respect to a local standard of rest (LSR velocity), as shown at the bottom of the figure, can be readily determined from the measured frequency offset of the observed line from the laboratory rest frequency of 28.533534 THz. In addition, a knowledge of the intrinsic stellar velocity from other types of observations can be used to infer the expansion velocity of the circumstellar gas with respect to the stellar core, as indicated at the top of the figure. (At a transition frequency of 28.5 THz ($\lambda=10.5\ \mu\text{m}$), each 20 MHz filter channel represents a Doppler-velocity shift of 0.2 km/sec.) Future stellar observations will emphasize studies of NH_3 in additional types of stars, and undoubtedly the detection of other interesting molecules in the stronger sources.

PROSPECTS

Although the 10 μm band is currently the most favorable spectral region for infrared heterodyne spectroscopy, if practical considerations of atmospheric transparency and front-end technology are considered, the 5 μm band also appears promising for stellar spectroscopy because of the astrophysically important bands of CO near 4.7 μm . HgCdTe photomixers function at this wavelength, and isotopic CO lasers appear an obvious choice for the LO. In addition, more complete spectral coverage at 5 μm may come from frequency-doubled CO_2 and N_2O lasers. Developmental work on a 5 μm heterodyne spectrometer is now starting at Berkeley. This emphasis on gas lasers as local oscillators is not meant to exclude other viable technologies, however. Following the work at other institutions, a smaller effort is now underway in our own laboratory to evaluate tunable diode lasers for laboratory spectroscopy and as possible local oscillators at frequencies unobtainable with gas lasers.

At a wavelength of 5 μm , Doppler widths are twice as broad as at 10 μm for a given velocity range. Thus, an IF bandwidth of 3 GHz will be needed to get the same velocity coverage provided by 1.5 GHz at 10 μm . To support the increased IF requirements both at 5 and 10 μm , another 64x20 MHz filterbank is now under construction. This will expand our simultaneous frequency coverage to 2.56 GHz and greatly improve the efficient use of observing time on a large telescope. The success of observations on a 1.5 m telescope shows us that in the future we can effectively use a 3 m instrument, such as the NASA 3 m infrared telescope facility in Hawaii, and thereby hope to gain a four-fold increase in signal-to-noise ratio on unresolved stellar sources. On such a large telescope, good spectroscopy of NH_3 and other circumstellar molecules will be possible on well over 100 stars.

REFERENCES

1. Freed, C.: "Designs and Experiments Relating to Stable Lasers", in Proc. Frequency Standards and Metrology Seminar (University Laval, Quebec, P.Q., Canada), Sept. 1, 1971, pp. 226-261.
2. Freed, C.; Bradley, L.C.; and O'Donnell, R.G.: "Absolute Frequencies of Lasing Transitions in Seven CO₂ Isotopic Species", IEEE J. Quant. Electr., (to be published Sept. 1980).
3. Skolnick, M.: "Use of Plasma Tube Impedance Variations to Frequency Stabilize a CO₂ Laser", IEEE J. Quant. Electr., QE-6, pp. 139-140, Feb. 1970.
4. Freed, Charles; and Javan, Ali: "Standing-Wave Saturation Resonances in the CO₂ 10.6 μ Transitions Observed in a Low-Pressure Room-Temperature Absorber Gas", Appl. Phys. Lett., 17, pp. 53-56, 15 July 1970.
5. Hillman, J.J.; Kostiuik, Theodor; Buhl, David; Faris, J.L.; Novaco, J.C.; and Mumma, M.J.: "Precision Measurements of NH₃ Spectral Lines Near 11 μ m Using the Infrared Heterodyne Technique", Optics Lett., 1, pp. 81-83, Sept. 1977.
6. Freund, S.M.; and Oka, Takeshi: "Infrared-Microwave Two-Photon Spectroscopy", Phys. Rev. A, 13, pp. 2178-2190, June 1976.
7. Ueda, Y.; and Shimoda, K.: "Infrared Laser Stark Spectroscopy", in Laser Spectroscopy II, ed. S. Haroche, (Lecture Notes in Physics, Vol. 43, pp. 186-197) Springer-Verlag: Berlin, Heidelberg, 1975.
8. Wells, J.S.; Petersen, F.R.; and Maki, A.G.: "Heterodyne Frequency Measurements with a Tunable Diode Laser--CO₂ Laser Spectrometer: Spectroscopic Reference Frequencies in the 9.5 μ m Band of Carbonyl Sulfide", Appl. Optics, 18, pp. 3567-3573, Nov. 1, 1979.
9. Johns, J.W.C.; McKellar, A.R.W.; and Riggins, M.: "Laser Magnetic Resonance Spectroscopy of the ν_3 Fundamental Band of HO₂ at 9.1 μ m" J. Chem. Phys., 68, pp. 3957-3966, May 1, 1978.
10. Johns, J.W.C.; McKellar, A.R.W.; and Riggins, M.: "Laser Magnetic Resonance Spectroscopy of the ν_2 Fundamental Band of HCO at 9.25 μ m," J. Chem. Phys. 67, pp. 2427-2435, Sept. 15, 1977.
11. Reid, John; and McKellar, A.R.W.: "Observation of the S₀(3) Pure Rotational Quadrupole Transition of H₂ with a Tunable Diode Laser", Phys. Rev. A., 18, pp. 224-227, July, 1978.
12. Spears, D.L.: "Planar HgCdTe Quadrantal Heterodyne Arrays with GHz Response at 10.6 μ m", Infrared Physics, 17, pp. 5-8, Jan. 1977.

13. Peterson, D.W.; Johnson, M.A.; and Betz, A.L.: "Infrared Heterodyne Spectroscopy of CO₂ on Mars", *Nature*, 250, pp. 128-130, July 12, 1974.
14. Betz, A.L.; McLaren, R.A.; Sutton, E.C.; and Johnson, M.A.: "Infrared Heterodyne Spectroscopy of CO₂ in the Atmosphere of Mars", *Icarus*, 30, pp. 650-662, 1977.
15. Betz, A.L.; Johnson, M.A.; McLaren, R.A.; and Sutton, E.C.: "Heterodyne Detection of CO₂ Emission Lines and Wind Velocities in the Atmosphere of Venus", *Astrophys.J.Lett.*, 208, pp. 141-144, Sept. 15, 1976.
16. Johnson, M.A.; Betz, A.L.; McLaren, R.A.; Sutton, E.C.; and Townes, C.H.: "Nonthermal 10 μ m CO₂ Emission Lines in the Atmospheres of Mars and Venus", *Astrophys.J.Lett.*, 208, pp. 145-148, Sept. 15, 1976.
17. Betz, A.L.; Sutton, E.C.; McLaren, R.A.; and McAlary, C.W.: "Laser Heterodyne Spectroscopy", in *Proc. Symposium on Planetary Atmospheres*, pp. 29-33, Royal Society of Canada, Ottawa, Ontario, Aug. 16-19, 1977.
18. Betz, A.L.; McLaren, R.A.; and Spears, D.L.: "NH₃ in IRC + 10216", *Astrophys.J.Lett.*, 229, pp. 97-100, April 15, 1979.
19. Betz, A.L.; and McLaren, R.A." "Infrared Heterodyne Spectroscopy of Circumstellar Molecules", in *Proc. IAU Symposium No. 87, Interstellar Molecules*, ed. B. H. Andrew; D. Reidel Co., Dordrecht (1980).
20. McLaren, R.A.; and Betz, A.L." "Infrared Observations of Circumstellar Ammonia in OH/IR Supergiants", *Astrophys. J. Lett.*, (to be published 1980).

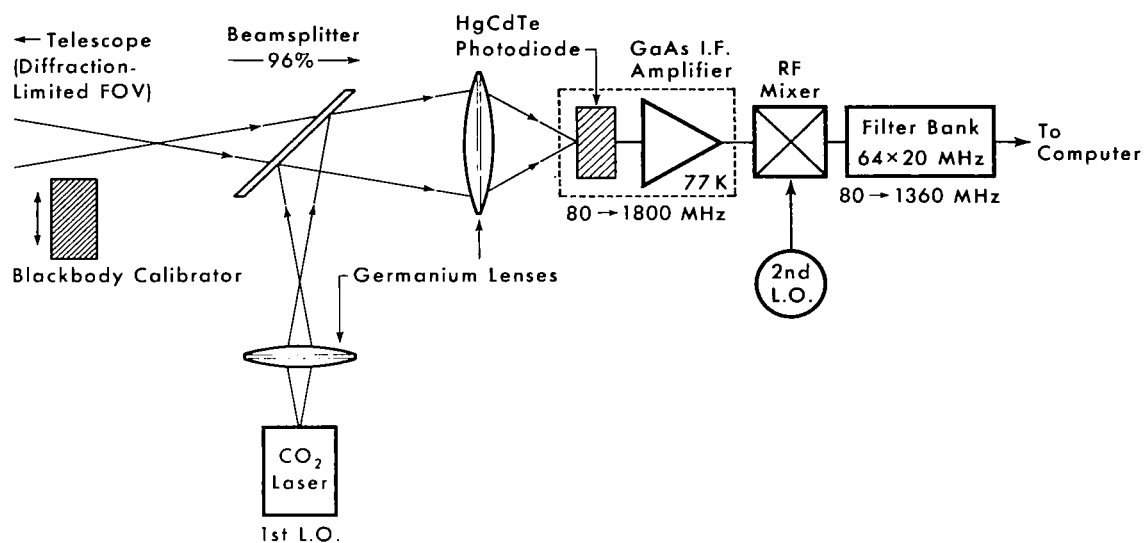


Figure 1.- Infrared heterodyne spectrometer.

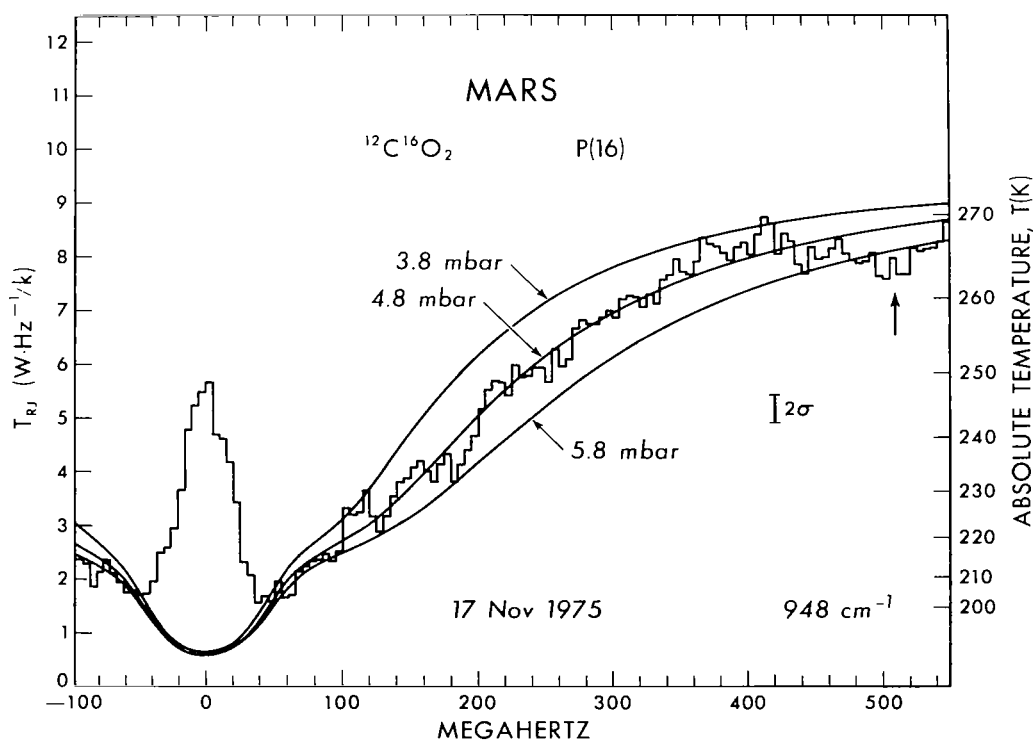


Figure 2.- Heterodyne measurements of CO₂ in Martian atmosphere.

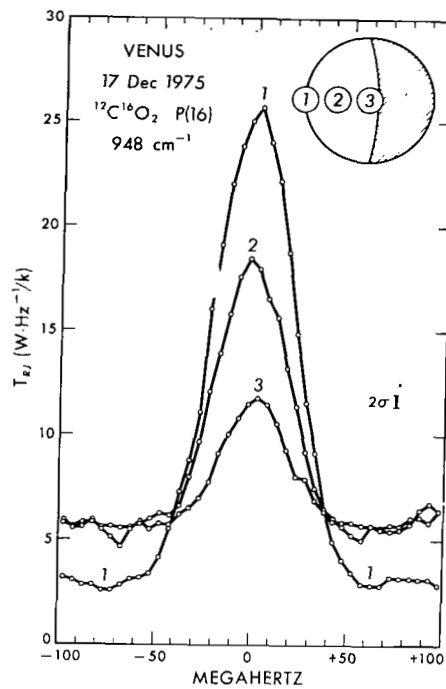


Figure 3.- Heterodyne measurements of CO_2 in Venusian atmosphere.

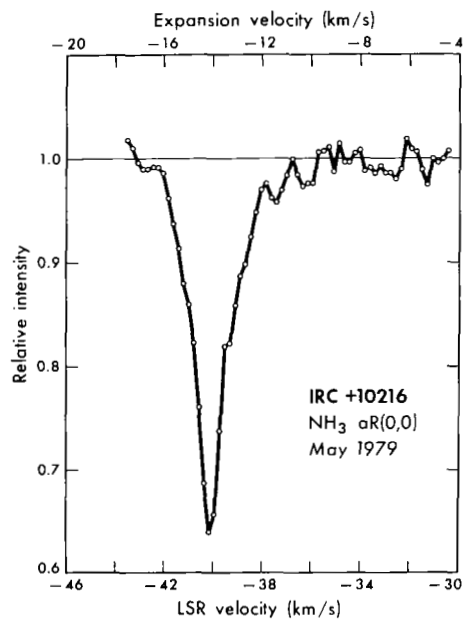


Figure 4.- Heterodyne measurements of NH_3 in star IRC + 10216 showing Doppler shift.

NEW PbSnTe HETEROJUNCTION LASER DIODE STRUCTURES WITH IMPROVED PERFORMANCE*

C. G. Fonstad, D. Kasemset, H. H. Hsieh[§], and S. Rotter
Department of Electrical Engineering and Computer Science
and Center for Materials Science and Engineering[†]
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

INTRODUCTION

In this article, we will summarize several of our recent advances in the state-of-the-art of lead tin telluride double heterojunction laser diodes, advances which make significant strides in increasing the operating temperatures of these devices and in controlling the modal quality and tunability of their output. CW operation to 120°K and pulsed operation to 166°K with single, lowest order transverse mode emission to in excess of four times threshold at 80°K have been achieved in buried stripe lasers fabricated by liquid phase epitaxy in the lattice-matched system, lead-tin telluride-lead telluride selenide [1,2]. At the same time, liquid phase epitaxy has been used to produce PbSnTe distributed feedback lasers with much broader continuous single mode tuning ranges than are available from Fabry-Perot lasers [3].

The physics and philosophy behind these advances is as important as the structures and performance of the specific devices embodying the advances, particularly since structures are continually being evolved and the performance continues to be improved. There is art in any science, but as we will demonstrate, there is a tremendous amount of science to be applied to Pb-salt tunable diode lasers, and where this is done, their performance can be predicted, tailored, and reproducibly controlled. Most importantly, their performance can be dramatically enhanced.

HIGH TEMPERATURE OPERATION

Achieving higher temperature operation of laser diodes requires both that the threshold current density be reduced at higher temperatures, and that the total threshold current be reduced. The former goal can be achieved by increasing the internal radiative efficiency of the lasing semiconductor, and by using a double heterojunction geometry to increase the efficiency with which stimulated recombination of the electrically injected carriers occurs. The latter goal requires development of a viable stripe geometry device achieving good lateral current confinement.

The double heterostructure was introduced to PbSnTe in the early 1970's [4,5] but the performance of the early devices was no better than that of diffused lasers. As we shall discuss first, this was due to the large degree of lattice mismatch inherent in these early devices which used PbTe substrates and confinement layers.

See footnotes at end of text, page 32.

Elimination of Lattice Mismatch

a. Role of Lattice Mismatch

Heterostructure devices make elegant use of semiconductor epi-layers with different energy gaps to achieve specific objectives, and laser diodes are an excellent example of such devices. Nonetheless, if these layers do not have nearly identical lattice parameters, a, an excessive number of defects will exist at the heterojunctions and in the bulk of the layers near them, and the desired performance will be degraded by these defects.

The surface morphology of mismatched epi-layers will also, in general, be much poorer than that of lattice matched layers. While this effect has not been quantified, it clearly has an important influence on the modal properties of the structures. The effect of mismatch created defects can be quantified in terms of an interface recombination velocity, s_v . This is in addition to the bulk lifetime, τ , which also can be affected by the mismatch. In the active region of a laser diode, which consists of a thin bulk layer of width d bounded by two heterojunctions, any excess minority carriers will see an effective lifetime which is lower than τ and decreases with increasing s_v and decreasing d , [6,7]

$$\tau_{\text{eff}}^{-1} \approx \tau^{-1} + 2s_v/d$$

By measuring τ_{eff} on many devices with different d , one can measure τ and s_v [7,8]. By also doing this measurement on samples with different amounts of mismatch, $\Delta a/a$, the dependence of s_v on mismatch can be found. In PbSnTe/PbTe, we find [9]

$$s_v = 2.9 \times 10^7 \Delta a/a \text{ cm/sec}$$

In a laser diode, the threshold current density is directly proportional to d/τ_{eff} , or in the limit of d going to zero, to $2s_v$ [10].

In traditional PbSnTe/PbTe DH laser diodes the mismatch can be very large and s_v can easily exceed 10^5 cm/sec. In a typical device, $\tau = 4$ nsec, $d = 1 \mu\text{m}$, and $s_v = 10^5$ cm/sec, so $\tau_{\text{eff}} = 0.8$ nsec and the threshold is five times higher than it would be if $s_v = 0$. Moreover, recent calculations have shown that the radiative lifetime should exceed 100 nsec [10] so the bulk lifetime, τ , is much lower than should be achievable with, in part, elimination of the mismatch.

b. Lattice Matched Systems

Except for the AlGaAs/GaAs combination, and to a lesser extent AlGaSb/GaSb, there are no lattice-matched binary/ternary combinations. In the III-V's, where it is to date only practical to grow binary substrates, it has thus been necessary to go to quaternary compounds such as InGaAsP to find a lattice-matched heteroepitaxy system. (InP is used as the substrate in this case.) In the IV-VI's we can grow bulk ternary substrates and it is possible to form lattice-matched heteroepitaxy systems from the ternaries. One such combination, which was first suggested by Walpole *et al.* [11], is PbSnTe and PbTeSe. Lead telluride selenide has the wider bandgap of this combination. This system can be used to cover the same bandgap range as PbSnTe. Another system covering a similar range is PbSSe (wider gap) and PbSnTe. A system covering the 4-6 μm range would be PbSTe (wider gap) and PbSSe. In all of these systems, the energy gap difference, and therefore the effectiveness of the heterostructure, de-

creases to zero as the compositions used approach the common binary, i.e. PbTe, PbSe, and PbS, respectively, in the three systems discussed. This is true in any lattice-matched system, of course, but is worth remembering, because it is important for short wavelength lasers.

c. Performance of Lattice-Matched Structures

Lead telluride selenide was chosen as the substrate material by Walpole et al. because of its wider bandgap, but it is difficult to work with (etch, polish, cleave, etc.), however, and perhaps for this reason, they saw little or no improvement in thresholds of their devices [12]. By using PbSnTe substrates we have overcome the problems of substrate preparation and end mirror cleaving, and have seen the anticipated threshold reductions in lattice-matched lasers[1].

The basic lattice-matched heterostructure begins with a p-type PbSnTe substrate upon which a p-type PbTeSe lower confinement layer, a 0.5-1.5 μm thick PbSnTe active layer, and an n-type PbTeSe upper confinement layer are grown. As with conventional lasers, ohmic contacts are then applied to both sides of the wafer, it is cut into bars with a wire saw, and individual lasers are cleaved from the bars. The 80°K threshold of these lasers is twenty times lower than that of comparable PbSnTe/PbTe lasers ($\lambda \approx 10 \mu\text{m}$) and pulsed operation has been achieved to 166°K [1]. In addition to these dramatic operating improvements, it is also found that the surface morphology, i.e. uniformity and smoothness, of the heterostructures, is much better than that of mismatched structures. In fact it typically equals that of the original substrate. Furthermore, the mode structure of even these broad area Fabry-Perot lasers is, qualitatively, cleaner than that of earlier structures.

Reduction of Total Current - Stripe Geometries

To optimize high temperature CW operation, a viable stripe geometry providing lateral current confinement is required. As will be discussed further below, it is also desirable for the structure to provide modest (as opposed to strong) lateral optical confinement to help achieve single mode output.

Two stripe techniques which have been used on Pb-salt lasers are the diffused stripe homojunction [13] and the selectively grown molecular beam epitaxy (MBE) mesa heterojunction stripe [14]. The first structure provides no optical confinement, while the latter provides too much. The latter also suffers because the thickness of the epi-layers tends to vary across the width of the stripe, and most importantly, the lasers often have very high excess currents and thus much higher threshold current densities than wide area lasers.

We initially investigated insulator-defined stripe contact and then etched mesa DH lasers. The former was unsuccessful because there is too much lateral current spreading through the low sheet resistance epi-layers and thus insufficient current confinement to yield any significant reduction in total threshold current. Interestingly, however, such lasers did show cleaner mode structure than broad area contact lasers. The etched mesa lasers showed excess junction leakage (discernable as a temperature invariant threshold at very low temperature [10]), which appeared to be occurring at the exposed junctions along the mesa walls. Anodization was partially successful in reducing this leakage but

the technique is not yet controllable enough to yield consistently repeatable results. The etched mesa has the inherent disadvantages of very strong lateral optical confinement and a very narrow top contact, too, so in light of the leakage problem, it seems most reasonable to abandon it at this time in favor of more generally attractive structures.

The buried mesa has recently received a great deal of attention from the III-V community because it offers strong carrier confinement and modest optical confinement in a relatively simple and easily controlled structure [15]. We have recently been successful in fabricating buried mesa DH lasers in the PbSnTe/PbTeSe system [2]. These first devices have low total threshold currents, 60-70 mA at 80°K, and emit over one hundred microwatts of power in a single, lowest order transverse mode at 80°K ($\approx 9 \mu\text{m}$). They have been operated CW to 120°K while mounted in a simple pressure contact test jig (from which they could be demounted). Modal properties of the devices will be discussed somewhat later.

The buried mesa lasers are fabricated by first growing a p-type PbTeSe layer and a PbSnTe active layer on a PbSnTe substrate. The wafer is then removed from the liquid phase epitaxy (LPE) system and photoresist stripes are patterned on the surface using standard photolithographic techniques. Mesas are next etched into the surface through the active layer and into the first epi-layer using the photoresist stripes as a mask. Then the surface is cleaned and the wafer is returned to the epitaxy system, where a final n-type (undoped) PbTeSe layer is grown over the entire surface of the wafer. Contacts are then applied to both sides of the wafer, the individual stripes are separated with a wire saw, and lasers are cleaved from the bars [2].

The present buried mesa lasers rely upon the fact that the current density across the PbTeSe-PbSnTe diode heterojunction in the stripe mesa should be much larger than that at the PbTeSe homojunction covering the rest of the device. Thus even though the area of the homojunction is 50 times as large as the heterojunction, most of the current should go through the stripe, i.e., the active region. In practice this was only partly successful, and the threshold indicated there was only partial current confinement. The reason for this is not yet fully understood, but it is perhaps due to too low a doping level in the upper n-type layer resulting in too large a sheet resistance in that layer. It may also be simply that what is needed is a structure wherein a truly blocking region is provided on either side of the mesa stripe. Such a structure will require the growth of additional layers and is currently under consideration.

Further Reduction in High Temperature Threshold

To achieve high temperature ($> 100^\circ\text{K}$) operation of Pb-salt lasers, not only must the low temperature threshold current density and total current be decreased as has been discussed, but the rate of increase in threshold with increasing temperature must also be reduced. This is an obvious statement but it is worth making because the high temperature behavior of Pb-salt lasers is dominated by a factor which has been ignored in the present discussion thus far,

and which is often neglected, i.e. the active layer doping level. The active layer must be very heavily doped in Pb-salt lasers [4c], on the order of 10^{18} cm^{-3} , because the high temperature threshold is directly proportional to the minority carrier density necessary to achieve population inversion in the active region [10], and this density decreases with increased doping levels [4c]. As the doping level is increased, however, at some point the minority carrier lifetime must surely decrease, and the free carrier absorption losses must increase; both effects will work to increase the threshold. The optimum active region doping level is, in fact, unknown, and this is perhaps one of the most important questions remaining to be answered. It should be given immediate attention. At the same time the control of doping levels and of dopants in the grown epi-layers must also be studied much more. It is not sufficient to know what the optimum doping level is because one must also be able to produce layers having the required doping profile.

As regards the buried mesa diodes mentioned above, the active region doping level was relatively low, and the rate of increase of their threshold with temperature was considerably more rapid than has been seen in broad area devices with more heavily doped active regions, and there is clearly room for improvement.

SINGLE MODE OUTPUT

A laser cavity like that found in a diode laser can have three types of families of modes: normal transverse modes, modes related to the portion of the cavity normal to the junction plane; lateral transverse modes related to the cavity in the junction plane but perpendicular to the direction of propagation; and longitudinal or axial modes arising from the long dimension of cavity in the junction plane. The normal transverse modes are controlled in the double heterojunction laser by making the active region sufficiently thin. The lateral transverse modes can be controlled by using a suitable stripe structure. The longitudinal modes are determined primarily by the spacing of the end mirrors in a Fabry-Perot laser, or the grating period in a distributed feedback laser. We will assume the active region can readily be made sufficiently thin so that the laser operates only in the lowest order normal lateral mode, and we will then look first at controlling the transverse lateral modes, and then the longitudinal modes.

Control of Transverse Modes

Mode control in the plane of the junction transverse to the direction of propagation can be achieved by providing a refractive index step in the plane either by passive, built-in means or by active, induced means. In the latter, there is transverse guiding or confinement only in the presence of an excitation current, i.e. gain guiding. The resulting index step is hard to control, however, and varies with the operating conditions.* A built-in index step is

* In an oxide defined stripe contact laser, gain guiding would be the primary transverse confinement mechanism.

more desirable and in fact should be large enough to dominate over any gain induced step that may potentially exist.

On the other hand, there is a distinct advantage in keeping the index step as small as possible. The smaller the index step, and thus, the weaker the guiding, the wider the guide can be and still support only the lowest order mode. If the index step is large, as is the case in an etched mesa structure, the stripe must be extremely narrow to ensure operation only in the lowest order transverse mode. In addition to being easier to fabricate, a wider laser will have a much higher output power than a very narrow one.

The transverse index step is greatly reduced in the buried mesa laser structure described earlier and those lasers do indeed operate in the lowest-order transverse mode up to four times threshold [2]. This has been confirmed by near and far field measurements of the emission, which showed one symmetrically shaped emission peak. Above four times threshold evidence of a second, higher order transverse mode family is seen although the emission remains substantially single mode to six times threshold.

The present buried mesa lasers have active regions $1.5\text{ }\mu\text{m}$ thick, and mesas $5\text{ }\mu\text{m}$ wide. If the active region is reduced to $0.5\text{ }\mu\text{m}$ thick, still wider mesas can be used. Nonetheless, the transverse index step is still larger than necessary to dominate over gain guiding and even wider stripes could be used if the guiding was weaker. A possible solution may be the incorporation of an additional, smaller index step within the buried mesa, possibly in the spirit of the channeled substrate planar (CSP) geometry lasers [16]. With such a combination our calculations indicate that stripe widths of 25 to 30 microns are practical. Fabricating such a device will require additional development of the LPE growth technology, however, and other solutions are also being explored.

Control of Longitudinal Modes

a. Role of Transverse and Lateral Modes

The importance of transverse mode control to the achievement of single mode operation in laser diodes has only recently been understood, but it is now generally accepted that a device that is constrained to operate only in one transverse mode (typically this is the lowest order mode) will oscillate over a broad range of drive currents in only one (longitudinal) mode. We can easily obtain lowest order normal transverse mode operation of a DH laser but unless a suitable stripe is used, there will be many lateral transverse modes, and there is a family of longitudinal (Fabry-Perot) modes associated with each transverse mode. Thus, while in a homogeneously broadened laser like a diode laser, one longitudinal mode in each family will dominate when operating well above threshold; there will still be many modes in the output because there are many transverse modes. Eliminating all but one transverse mode eliminates all but one family of Fabry-Perot modes, and leads to single mode output. Consequently, the key to obtaining a single longitudinal mode output, i.e. truly single mode output, lies in achieving operation in only the lowest order transverse mode.

b. Distributed Feedback

As was just discussed, one longitudinal mode in each transverse mode family will dominate and that mode is the one nearest the peak of the gain curve. As the junction temperature of the laser is changed, i.e., as it is tuned, the gain curve shifts and thus the preferred longitudinal mode may change and mode hopping may occur. This hopping will be frequent in the case of a Fabry-Perot (FP) laser where the longitudinal modes are closely spaced.* The spacing of the cavity modes can be increased, and thus the incidence of mode hopping decreased, by shortening the laser, but this also increases the threshold (which varies as $1/L$) and is thus a clear compromise solution.

Another solution is to replace the Fabry-Perot end-mirror cavity by a distributed cavity formed by making a weak periodic variation in refractive index within the active region in the direction of the propagation of the light. Such a distributed feedback (DFB) cavity also has a family of closely spaced modes but the modes away from the two nearest the Bragg frequency of the periodic variation or "grating" have much higher thresholds and there is strong discrimination against them. Thus, a DFB laser will lase in a single mode, the near-Bragg mode nearest the gain peak, over a much wider tuning range than an FP laser will before mode hopping occurs. On the other hand, the DFB laser will only lase at those temperatures for which the gain curve overlaps the Bragg frequency, whereas there are always modes that overlap the gain curve in an FP laser.

We have successfully fabricated PbSnTe DFB lasers by liquid phase epitaxy and observed unusually clean, single mode emission spectra continuously tunable over a range in excess of 20 cm^{-1} ($0.33 \text{ }\mu\text{m}$) centered about 780 cm^{-1} ($12.8 \text{ }\mu\text{m}$) at an average rate of $1.2 \text{ cm}^{-1}/^\circ\text{K}$ from 9°K to 26°K .

The DFB lasers were produced by LPE growth of PbSnTe ($1.5 \text{ }\mu\text{m}$ thick) and PbTe ($0.5 \text{ }\mu\text{m}$) layers directly on a $0.79 \text{ }\mu\text{m}$ period corrugation ion milled into a PbTe substrate, and had a $25 \text{ }\mu\text{m}$ wide insulator defined stripe contact extending over $400 \text{ }\mu\text{m}$ of their overall length of $710 \text{ }\mu\text{m}$. Fabry-Perot modes were thus eliminated by providing an unpumped absorbing region on one end of the device; the end of the absorbing region was also saw-cut at an angle of approximately 30° to the stripe axis to further reduce FP reflections.

These DFB lasers are unique in that they are fabricated by liquid phase epitaxy and in that the feedback corrugation is placed within the active region of the device and is produced before, rather than after, the growth of the active layer. The tuning rate and range, and the cleanliness of the mode spectrum and the degree of single mode operation of these devices are superior to previously reported PbSnTe DFB lasers [17]. The full operating range of these devices is in fact still unknown because they were already operating at the lowest temperature and even broader tuning ranges may be possible.

These DFB lasers had moderate threshold 3.6 kA/cm^2 (pulsed). While the

*The mode spacing is λ^2/nL , where λ is the lasing wavelength, L the cavity length, and n the index of refraction.

output power vs. current relationship is not completely free of kinks, substantially single mode operation can be obtained over the entire operating range and to over 10 times threshold. Weaker modes can at times be seen in the spectrum, however, which indicates that the transverse modes are not adequately suppressed by the stripe contact.

While these DFB lasers clearly demonstrate the merit of this structure, they were not fabricated in the lattice matched system described earlier nor was the mesa stripe employed and thus there remains much room for improvement. On the other hand, the technology for fabricating micron period corrugations and the stripe mesa technology are both sufficiently new that it is felt that the immediate emphasis should be placed on fully developing those technologies, after which their combination to produce optimal DFB lasers will be relatively straightforward.

CONCLUSIONS AND FUTURE PRIORITIES

The recent developments described above demonstrate the value of using sophisticated device geometries to improve the performance of Pb-salt laser diodes. It should be clear, furthermore, that only through the use of heterostructures and stripe geometry devices with built-in index steps can the goals of high temperature CW operation and single mode, low noise emission be reproducibly achieved. To fabricate such structures requires the use of epitaxial techniques and of more complex processing sequences than are commonly used with Pb-salt devices at present, but the gains to be realized are tremendous.

With each new solution to yesterday's problems, and with each device refinement, device performance is improved but new limitations are encountered which, if overcome, promise still better performance. With Pb-salt lasers, the most important immediate goal remains the perfection of the stripe geometry, single mode device in a lattice-matched system. While nothing has been said in this presentation about the impact of this on emission noise, it is expected that such devices will have much lower noise, and determining if this is indeed the case will be the next question to answer. Beyond that the most important basic issues before us appear to be controlling and optimizing doping profiles in devices; determining and minimizing the role of defects and impurities in limiting the internal radiative recombination efficiency; understanding the limitations on laser output power at high pumping levels; and understanding sources of, and solutions for, noise in the output of these lasers.

REFERENCES

1. Kasemset, D.; and Fonstad, C. G.: Lattice-Matched $\text{Pb}_{1-x}\text{Sn}_x\text{Te}/\text{PbTe}_{1-y}\text{Se}_y$ DH Laser Diodes Operating to 166K. Proceedings of the 1979 International Electron Devices Meeting, IEEE, 1979. pp. 130-132.
2. Kasemset, D.; Rotter, S.; and Fonstad, C. G.: $\text{Pb}_{1-x}\text{Sn}_x\text{Te}/\text{PbTe}_{1-y}\text{Se}_y$ Lattice-Matched Buried Heterostructure Lasers with CW Single Mode Output. IEEE Electron Device Letters, vol. EDL-1, May 1980, pp. 75-78.
3. Hsieh, H. H.; and Fonstad, C. G.: LPE-Grown PbSnTe Distributed Feedback Laser Diodes with Broad Continuous Tuning Range in Single Mode Output. Proceedings of the 1979 International Electron Devices Meeting, IEEE, 1979. pp. 126-129.
- 4a. Tomasetta, L. R.; and Fonstad, C. G.: Liquid Phase Epitaxial Growth of Laser Heterostructures in $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$. Appl. Phys. Lett., vol. 24, 1974, pp. 567-570.
- 4b. Tomasetta, L. R.; and Fonstad, C. G.: Threshold Reduction in $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Laser Diodes through the use of Double Heterojunction Geometries. Appl. Phys. Lett., vol. 24, 1974, pp. 440-442.
- 4c. Tomasetta, L. R.; and Fonstad, C. G.: Lead-Tin Telluride Double Heterojunction Laser Diodes: Theory and Experiment. IEEE J. Quant. Electron., vol. QE-11, 1975, pp. 384-390.
- 5a. Walpole, J. N.; Calawa, A. R.; Ralston, R.; Harmon, T.; and McVittie, J.: Single Heterojunction $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Diode Lasers. Appl. Phys. Lett., vol. 23, 1973, pp. 620-622.
- 5b. Groves, S. H.; Nill, K. W.; and Strauss, A. J.: Double Heterostructure $\text{Pb}_{1-x}\text{Sn}_x\text{Te}-\text{PbTe}$ Lasers with CW Operation at 77K. Appl. Phys. Lett., vol. 25, 1974, pp. 331-333.
6. Armiento, C. A.; and Fonstad, C. G.: An Analysis of the Effects of Interface Recombination on the Transient Response of Double Heterojunction Devices. IEEE J. Quant. Electron., vol. QE-13, 1977, pp. 783-791.
7. Fonstad, C. G.; and Armiento, C. A.: Approximating the Transient Response of Double Heterojunction Devices. J. Appl. Phys., vol. 49, 1978, pp. 2435-2438.
8. Kasemset, D.; and Fonstad, C. G.: Measurement of Interface Recombination Velocity in $(\text{Pb},\text{Sn})\text{Te}/\text{PbTe}$ Double Heterostructure Laser Diodes. Appl. Phys. Lett., vol. 34, 1979, pp. 432-434.
9. Kasemset, D.; and Fonstad, C. G.: Reduction of Interface Recombination Velocity with Decreasing Lattice Parameter Mismatch in PbSnTe Heterojunctions. J. Appl. Phys., vol. 50, 1979, pp. 5028-5029.

10. Kasemset, D.; and Fonstad, C. G.: Minority Carrier Lifetimes and Lasing Thresholds of PbSnTe Heterostructure Lasers. IEEE J. Quant. Electron., vol. QE-15, 1979, pp. 1266-1270.
11. Walpole, J. N.; Groves, S. H.; Calawa, A. R.; and Harman, T. C.: Lattice Misfit Dislocations in Heteroepitaxial $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$. Solid State Research Report, M.I.T. Lincoln Laboratory, No. 4, 1974, p. 13.
12. Walpole, J. N.: private communication.
13. Ralston, R. W.; Melngailis, I.; Calawa, A. R.; and Lindley, W. T.: Stripe Geometry $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Diode Lasers. IEEE J. Quant. Electron., vol. QE-9, 1973, pp. 350-356.
14. Walpole, J. N.; Calawa, A. R.; Harman, T. C.; and Groves, S. H.: Double Heterostructure PbSnTe Lasers Grown by Molecular Beam Epitaxy with CW Operation up to 114K. Appl. Phys. Lett., vol. 28, 1976, pp. 552-554.
- 15a. Tsukuda, T.: GaAs- $\text{Ga}_{1-x}\text{Al}_x\text{As}$ Buried Heterostructure Injection Lasers. J. Appl. Phys., vol. 45, 1974, pp. 4899-4906.
- 15b. Kishino, K.; Suematsu, Y.; Takahashi, Y.; Tanbun-ek, T.; and Itaya, Y.: Fabrication and Lasing Properties of Mesa Substrate Buried Heterostructure GaInAsP/InP Lasers at 1.3 μm Wavelength. IEEE J. Quant. Electron., vol. QE-16, 1980, pp. 160-164.
16. Aiki, K.; Nakamura, M.; Kuroda, T.; Umeda, J.; Ito, R.; Chinone, N.; and Maeda, M.: Transverse Mode Stabilized $\text{Al}_x\text{Ga}_{1-x}\text{As}$ Injection Lasers with Channeled-Substrate-Planar Structure. IEEE J. Quant. Electron., vol. QE-14, 1978, pp. 89-94.
- 17a. Walpole, J. N.; Calawa, A. R.; Chinn, S. R.; Groves, S. H.; and Harman, T. C.: Distributed Feedback $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Double Heterostructure Lasers. Appl. Phys. Lett., vol. 29, 1976, pp. 307-309.
- 17b. Walpole, N. J.; Calawa, A. R.; Chinn, S. R.; Groves, S. H.; and Harman, T. C.: CW Operation of Distributed Feedback $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Lasers. Appl. Phys. Lett., vol. 30, 1977, pp. 524-526.

*This work sponsored by NASA Grant NSG 1413 and, through the MIT Center for Materials Science and Engineering, by NSF Grant No. DMR-76-80895.

§Present address: IBM, Yorktown Heights, N.Y.

†The Center is an NSF Materials Research Laboratory supported in part by NSF Grant No. DMR-76-80895.

ADVANCES IN TUNABLE DIODE-LASER TECHNOLOGY

Wayne Lo
General Motors Research Laboratories

SUMMARY

Three major areas of research have been investigated. They are: the improvement of long-term reliability, the purification of mode properties, and the achievement of higher-temperature operation. In reliability studies we observed a slow increase in contact resistance during room temperature storage for lasers fabricated with In-Au or In-Pt contacts. This increase is actually caused by the diffusion of In into the surface layer of laser crystals. By using a three-layered structure of In-Au-Pt or In-Pt-Au, we have reduced this mode of degradation. In characterizing the mode properties, we found that the lasers emit in a highly localized, filamentary manner. For wide-stripe lasers the emission occurs near the corners of the junction. In order to achieve single-mode operation, stripe widths on the order of 8-10 μm are needed. We also found that room temperature electroluminescence is possible near 4.6 μm .

INTRODUCTION

In recent years the technology of fabricating lead-salt diode lasers has advanced rapidly. Progress has been stimulated by the demand for more reliable and well behaved devices for certain applications. For instance, deleterious changes in electrical and optical properties have been overcome.^{1,2} This was achieved by improving the crystal growth technique, the diffusion process, and above all the contacting method. We have found that contact reliability is the dominant factor in determining the lifetime of lead-salt diode lasers. Besides long term reliability, better emission characteristics are essential for most applications. We have characterized stripe-geometry $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ diode lasers with data on near-field measurements. In this paper, the technology needed to fabricate simple single-mode, stripe-geometry lasers will be discussed.

Higher temperature operation is another desirable feature of well behaved lasers. Processes leading to the fabrication of room temperature $\text{PbS}_{1-x}\text{Se}_x$ light emitting diodes will be discussed.

Contact Reliability

For cw operation at temperatures approaching 40 K, we found that a contact resistance of 10^{-4} ohm-cm^2 or less is necessary. This is attributed

to the low thermal conductivity of lead salt materials (typically 0.07 W/cm-K versus 3.0 W/cm-K for GaAs at 77 K).⁵ Since lead-salt crystals have very high electron affinities, low resistance and ohmic contact to the n-type side of diodes can be obtained easily by using either Au, In or In-Au combinations. It is the p-type contact that normally causes problems.

For the p-type side In-Au,⁴ In-Pt⁵ and In-Pt-Au⁶ structures have been reported as producing low resistance, ohmic contacts. More recently only In-Au combinations have been used.^{7,8} We are not aware of any systematic studies on the long-term reliability of In-Au contacts. However, we found that they are not stable in our devices. During storage at room temperature over periods that range from days to months, the resistance of In-Au contacts can increase. Such increases are typically accompanied by increases in cw threshold current densities and laser emission frequencies, while at the same time laser output powers tend to decrease.

We found that the increase in contact resistance is caused by the diffusion of In into the surface layer of the p-type side. Figure 1 shows an electron microprobe analysis of In and Pb concentration profiles near the metal-semiconductor interface before and after degradation. The samples were prepared with In-Au as the p-type contact. The degradation was accelerated by heating the sample to 100°C for one hour. This test condition is equivalent to room temperature storage of approximately one week. Both the degraded and the control samples were angle-lapped at a 3° inclination to provide a twenty-fold improvement in spatial resolution normal to the interfaces.

The typical electron beam penetration depth is about 1 to 2 μm . The X-rays generated by the E-beam (the signal strength shown in Fig. 1) are absorbed partially by the contact metals before being detected. Au and Pt are good absorbers for the X-rays. We calculated that 70% of the X-rays generated in the crystal are absorbed by Au in a distance of 0.2 μm . This explains the sharp change in the Pb X-ray signal near the interface (and hence the good accuracy in identifying the interface). From Fig. 1 it is also noted that for a degraded sample, traces of In were found in the first few μm of the crystal below the contact. Since In is a donor in lead-salt materials, a reduction in hole carrier concentration near the surface is expected. We think that this is the reason for the increase in contact resistance.⁹

Further study revealed that Au or Pt alone cannot form a barrier against In penetration, but a combination of Pt-Au does.^{1,2} It was also noticed that the reliability of the lasers depends on the thickness of the Pt and Au layers. We found that both layers have to be at least 0.2 μm thick in order to form a barrier against In diffusion. After depositing the Au and Pt layers on the p-type side of the diode a 10 μm layer of In is added. The n-type contact consists of a 0.2 μm layer of Au plus 5 μm of In. The p-type side of the crystal surface was purposely oxidized by exposure to air before evaporating the first layer of Au. This increases the hole carrier concentration near the surface and tends to stabilize the contact resistance.

Over a period of ten months, groups of $\text{PbS}_{1-x}\text{Se}_x$ and $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ lasers (made with three-layer contacts on the p-type side) have been tested for thermal cycling, room temperature storage and cw operation. For up to sixty thermal cycles and five hundred hours of cw operation, no significant changes in contact resistance, threshold current density or optical properties were observed. The test results are summarized in Table 1. This is a significant improvement over lasers that have been fabricated in the past, whose performance could degrade in a matter of a few days at room temperature.

Mode Control

Diode Lasers with both contact stripes and diffused junction stripes have been studied. Since lead-salt materials generally have very high electrical conductivity, lasers with contact stripes will encounter severe current spreading problems.¹⁰ For this reason, the lasers described here were made with junction stripes by an impurity diffusion process. Both Sb and Cd were used as diffusion sources,^{11,12} with MgF_2 and SiO_2 as the diffusion masks.

In lasers with 125 μm wide stripes, the emission originated from the corners of the diffused planar junctions. The near field pattern exhibited two peaks with a separation approximately equal to the width of the stripe (Fig. 2). This is consistent with the tuning curve that was observed with a monochromator. By adjusting the position of a collection lens in front of the laser, each filament could be imaged separately on the entrance slit of the monochromator. As shown in Fig. 3 the two filaments produced two separate, but parallel, tuning curves. The frequency difference was approximately constant at 5.3 cm^{-1} . This difference could be produced by a 0.1% variation in the composition of the $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ across the 125 μm stripe. When the detector was scanned along the direction perpendicular to the p-n junction plane and near one corner of the stripe, the near field pattern showed only one peak. This suggested that the photon confinement perpendicular to the junction was quite good for these impurity-diffused junctions.¹³ Proper confinement parallel to the junction was then expected to produce single-filament operation. We estimate that a stripe width on the order of 10 μm is needed to achieve single-mode operation.

From Fig. 2 it is noted that the spot size of the filaments is quite broad. Since a simple BaF_2 lens was used in this experiment, the resolution is set by aberration. By using the third-order approximation of a perfectly centered thin-lens aberration theory, we estimated that the image of an infinitely distant axial point source would be approximately 44 μm . This is in good agreement with the spot size observed in Fig. 2. In order to see any fine structure, a diffraction-limited lens has to be used.

OPERATING TEMPERATURE

In an effort to predict the highest temperature of operation for lasers at long wavelengths, we have made an empirical plot as shown in Fig. 4. From this plot, it is seen that the longest wavelength for a room temperature semiconductor laser is $2.5\text{ }\mu\text{m}$. Although laser sources are desirable for many applications, room temperature, long wavelength incoherent light sources are of interest for low resolution applications, as well as for low loss optical fiber communication systems.^{14,15} The narrow band gap, lead-salt materials are suitable for fabricating emitters in the $3\text{--}16\text{ }\mu\text{m}$ spectral range. Photoluminescence¹⁶ and cathodoluminescence¹⁷ have been observed at room temperature, but not electroluminescence. We report here the first recorded electroluminescence from lead-sulfide-selenide at room temperature.

The diodes were fabricated from high quality lead-sulfide-selenide single crystals grown from the vapor phase. A two-step diffusion-annealing process produced a graded carrier concentration region¹⁸ for photon confinement transverse to the junction. No attempt was made to confine the carriers or photons along the junction plane. These results represent a further improvement in the two-step diffusion-annealing process which produced low temperature diode lasers in lead-sulfide-selenide.¹⁸

High temperature spontaneous emission spectra were recorded as shown in Fig. 5. Notice that the 330 ppm atmospheric CO_2 absorption band near $4.2\text{ }\mu\text{m}$ appears in the 260 K and 300 K spectra (Figs. 5(e) and 5(f)).

The power output for this (laser) diode geometry is low, on the order of a few hundred nW. This is attributed to the high index of refraction (4.6) for lead-sulfide-selenide, which produces substantial internal reflection at the semiconductor-air interface. The critical angle for total internal reflection (i.e., the largest angle which incident radiation can subtend with the normal and still exit the crystal) is 12° . A large fraction of the randomly oriented junction radiation is totally reflected back into the semiconductor and absorbed. By using a hemispherical structure and extracting light from the top surface rather than the side, a factor of 1000 improvement in power is possible. This will allow these diodes to emit $200\text{ }\mu\text{W}$ at room temperature.

CONCLUDING REMARKS

In conclusion, we found that the slow degradation of contact resistance for lead-salt diode lasers fabricated with In-Au contacts is due to the migration of In into the surface layer of the p-type side of the lasers. By using combinations of Au-Pt or Pt-Au as barriers to prevent In migration, we have substantially increased the lifetime of these lasers. For laser mode control, we found that narrow stripe ($\sim 10\text{ }\mu\text{m}$) diode lasers are needed for single-mode operation. Finally, for high temperature operation, we have shown that although lasing action at room temperature is possible only up to $2.5\text{ }\mu\text{m}$, spontaneous emission at room temperature has been observed near $4.6\text{ }\mu\text{m}$.

REFERENCES

1. Lo, W.: The Elimination of Contact Degradation in Lead-Salt Diode Lasers. Jap. J. Appl. Phys., Vol. 18, Suppl. 18-1, 1979, pp. 367-369.
2. Lo, W.; and Gifford, F. E.: Contact Reliability Studies on Lead-Salt Diode Lasers. J. of Electrochemical Society, June 1980.
3. American Institute of Physics Handbook. McGraw Hill, 1975, pp. 4-156.
4. Ralston, R. W.; Melngailis, I.; Calawa, A. R.; and Lindi, W. T.: Stripe-Geometry $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Diode Lasers. IEEE, J. Quant. Electron., Vol. QE-9, 1973, pp. 350-356.
5. Antcliff, G. A.; and Parker, S. G.: Characteristics of Tunable $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Junction Lasers in the 8-12 μm Region. J. Appl. Phys., Vol. 44, 1973, pp. 4145-4160.
6. Ralston, R. W.; Walpole, J. N.; Calawa, A. R.; Harman, T. C.; and McVittie, J. P.: High cw Output Power in Stripe Geometry PbS Diode Lasers. J. Appl. Phys., Vol. 45, 1974, pp. 1323-1325.
7. Walpole, J. N.; Calawa, A. R.; Harman, T. C.; and Grove, S. H.: Double Heterostructure $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Lasers Grown by Molecular-Beam Epitaxy. Appl. Phys. Lett., Vol. 28, 1976, pp. 552-554.
8. Linden, K. J.; Nill, K. W.; and Butler, J. F.: Single Heterostructure Lasers of $\text{PbS}_{1-x}\text{Se}_x$ and $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ with Wide Tunability. IEEE J. Quant. Electron., Vol. QE-13, No. 8, 1977, pp. 720-725.
9. Lo, W.: Te-Rich Growth and Laser Fabrication of $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ with $0.06 < x < 0.08$. J. Electron. Mater., Vol. 6, 1977, pp. 39-48.
10. Dumke, W. P.: Current Threshold in Stripe-Contact Injection Lasers. Solid State Electron., Vol. 16, 1973, pp. 1279-1281.
11. Lo, W.; Montgomery, G. P., Jr.; and Swets, D. E.: Ingot-Nucleated $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Diode Lasers. J. Appl. Phys., Vol. 47, 1976, pp. 267-271.
12. Lo, W.: Cd-Diffused $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Diode Lasers with High Output Power. Appl. Phys. Lett., Vol. 28, 1976, pp. 154-156.
13. Lo, W.: Homojunction $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ Diode Lasers with Increased Frequency Tuning Range. IEEE J. Quant. Electron., Vol. QE-13, 1977, pp. 591-595.
14. Pinnow, D. A.; Gentile, A. L.; Standlee, A. G.; Timper, A. J.; and Hobrock, L. M.: Polycrystalline Fiber Optical Waveguide for Infrared Transmission. Appl. Phys. Lett., Vol. 33, 1978, pp. 28-29.

15. VanUitert, L. G.; and Wemple, S. H.: ZnCl_2 Glass: A Potential Ultralow-Loss Optical Fiber Material. Appl. Phys. Lett., Vol. 33, 1978, pp. 57-59.
16. Galeski, F.; Drozd, L.; Lebedeva, L. Y.; Ten, V. P.; and Yunovich, A. E.: Stimulated Emission From Thin PbSe Films at Room Temperature. Sov. Phys. Semicond., Vol. 11, 1977, pp. 327-328.
17. Kurbatov, L. N.; Britov, A. D.; and Karavaev, S. M.: Cathodoluminescence of Lead-Tin Selenide. Sov. Phys. Semicond., Vol. 9, 1976, pp. 1045-1046.
18. Lo, W.; and Swets, D. E.: Diffused Homojunction Lead-Sulfide-Selenide Diodes with 140 K Laser Operation. Appl. Phys. Lett., Vol. 33, 1979, pp. 938-940.

Table 1. Laser test results
(Ten-month test period)

		R_c (Ohm-cm ²) $\times 10^5$ Contact Resistance		J_{th} (A/cm ²) Threshold Current density		P(mW) Total Output Power at I=1 Amp, T=20K		Remarks
$PbS_{0.82}Se_{0.18}$		Initial	Final	Initial	Final	Initial	Final	
	#1	1.10	1.18	260	280	0.85	0.80	•500 hours c.w. operation •Thermal cycling 26 times
	#2	0.87	0.91	365	380	3.2	2.6	•Thermal cycling 45 times
$Pb_{0.86}Sn_{0.14}Te$	#1	0.74	0.80	95	105	0.65	0.58	•300 hours c.w. operation •Thermal cycling 25 times
	#2	0.65	0.73	125	135	0.90	0.86	•420 hours c.w. operation •Thermal cycling 33 times

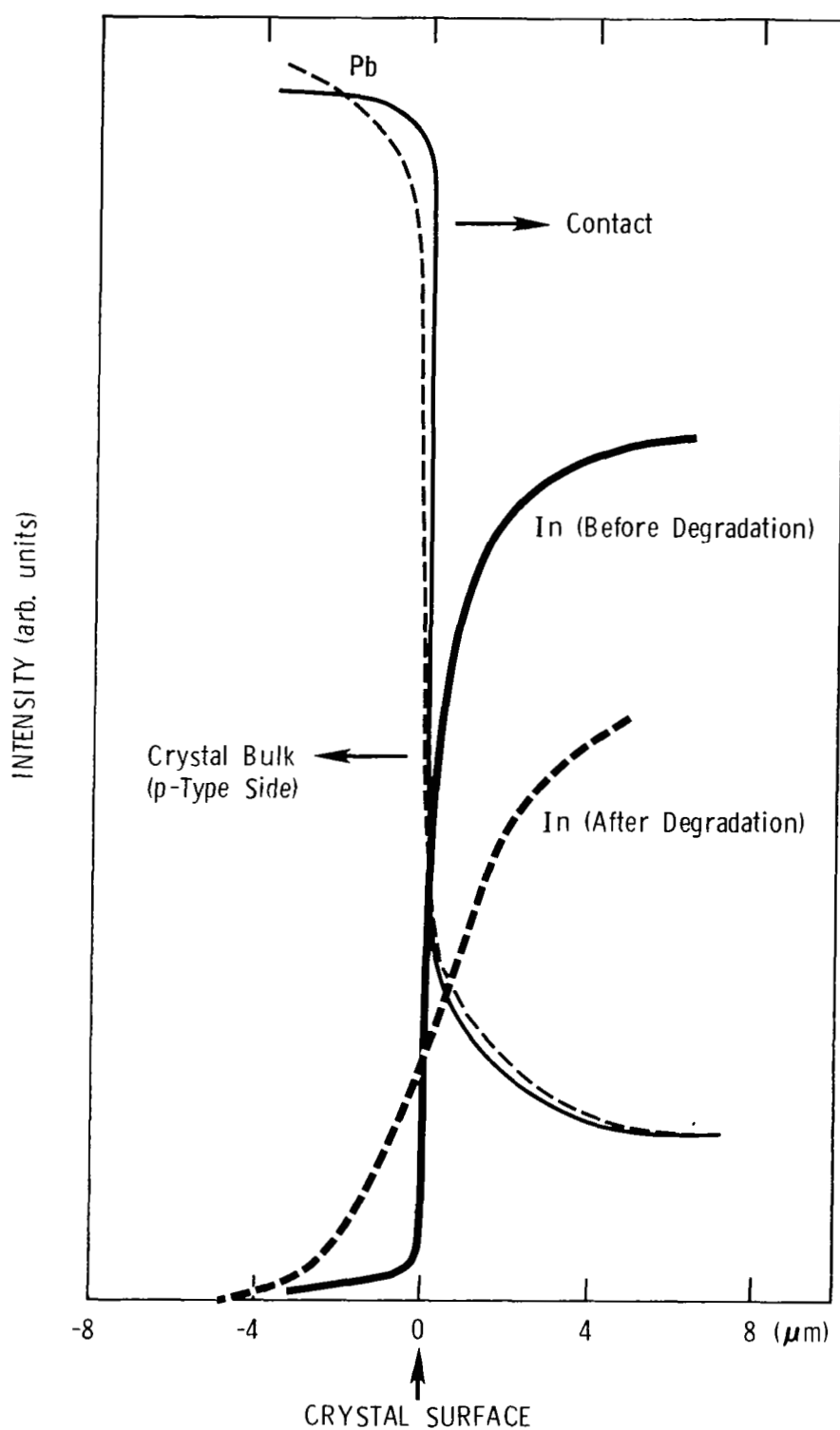


Figure 1.- Electron microprobe analysis of a crystal-contact interface for $\text{PbS}_{0.82}\text{Se}_{0.18}$.

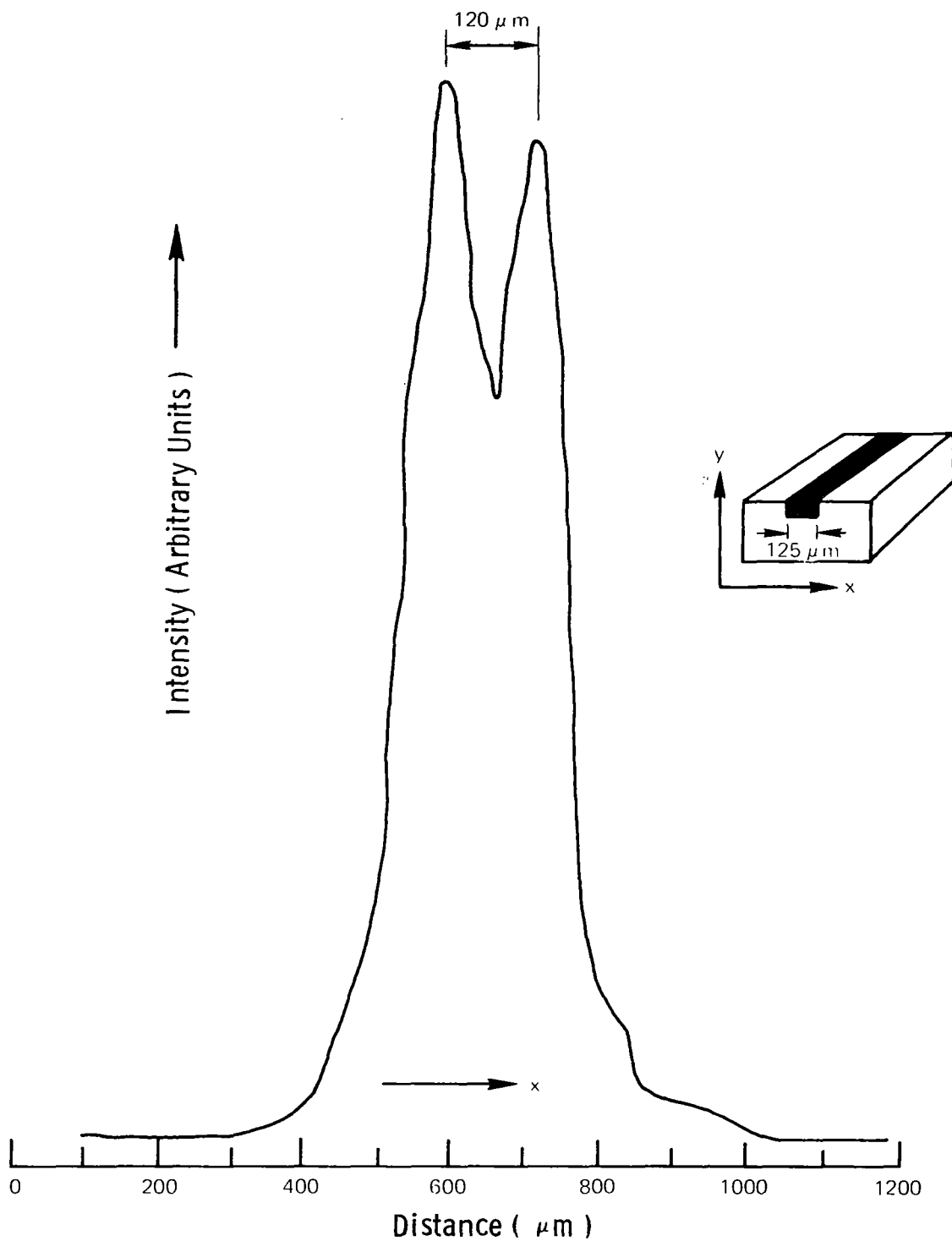


Figure 2.- Near field pattern in the plane parallel to the junction for a $125\ \mu\text{m}$ stripe laser.

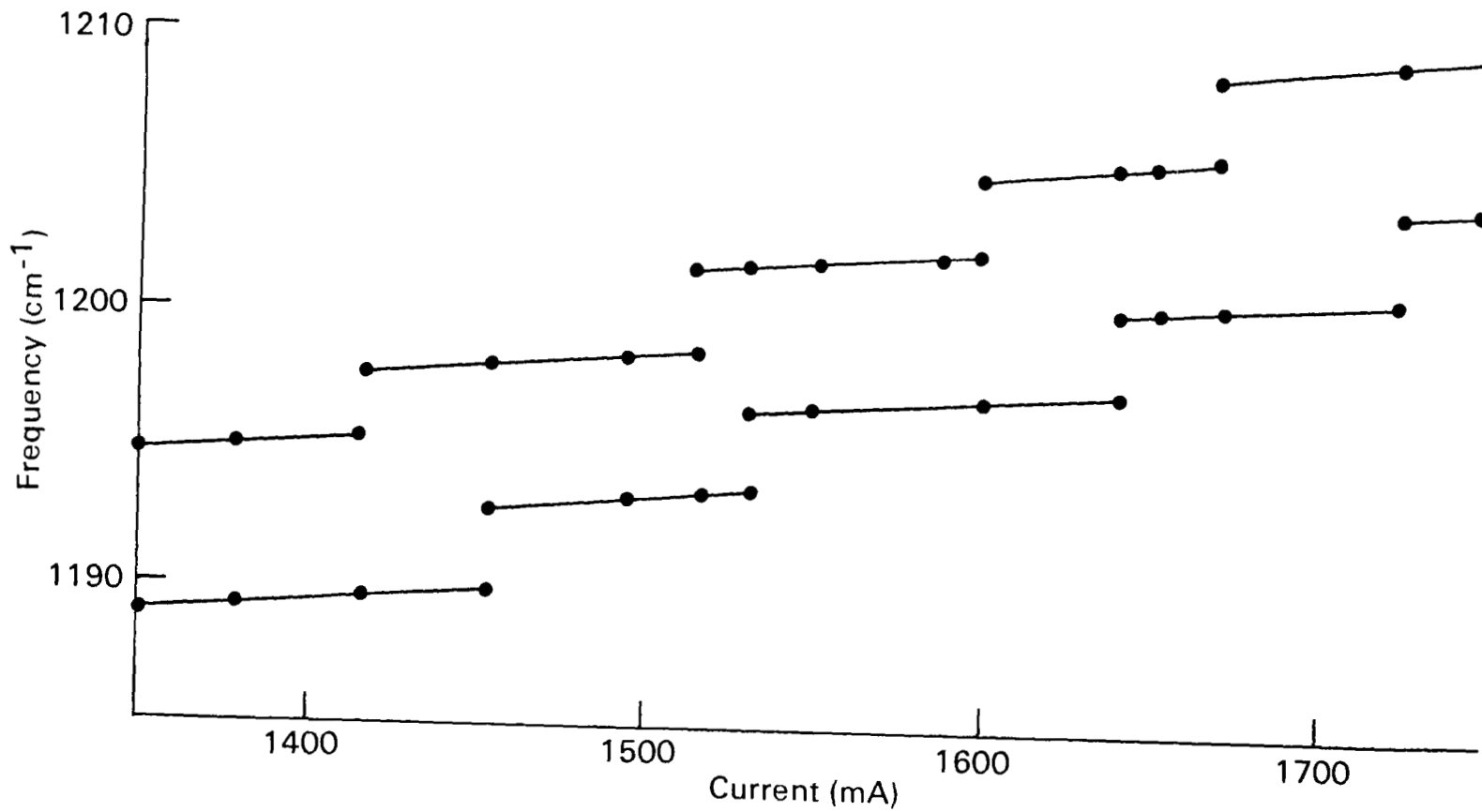


Figure 3.- Tuning curve for a 125 μm stripe laser.

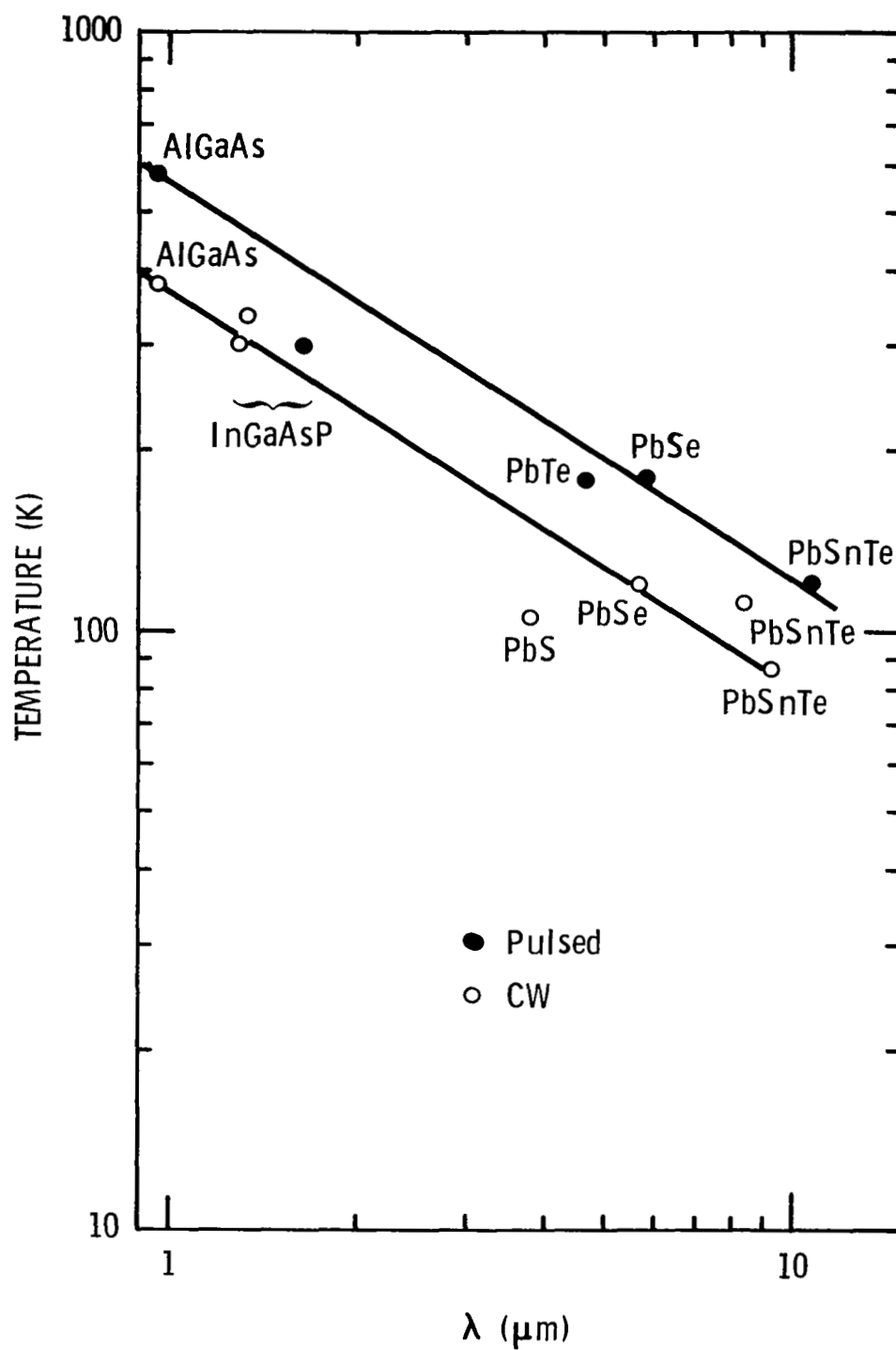


Figure 4.- Empirical plot of the highest observed operating temperatures for semiconductor diode lasers at different wavelengths.

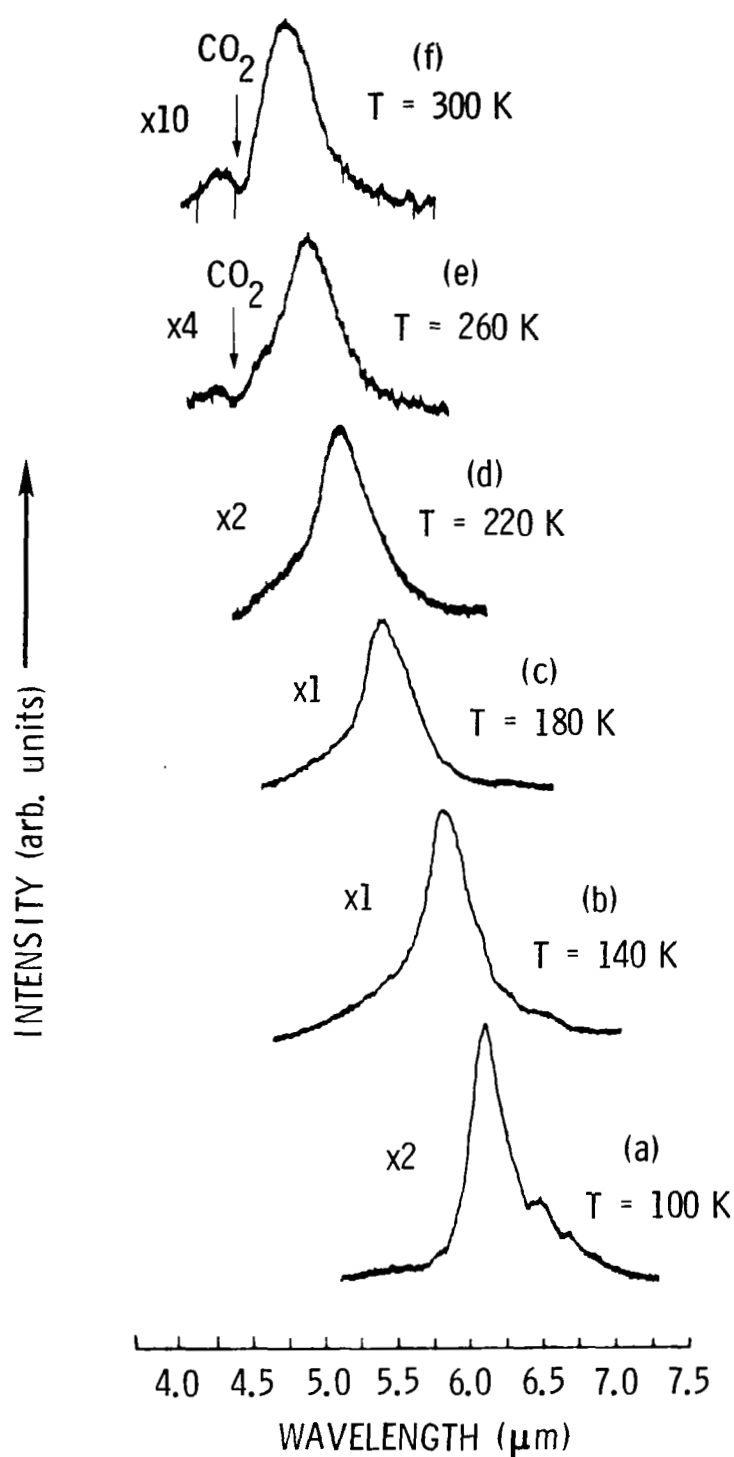


Figure 5.- Spontaneous emission spectra at different temperatures. Also shown is the 330 ppm atmospheric CO_2 absorption band near 4.2 μm in the 260K and 300K spectra.

LONG WAVELENGTH PbSnTe LASERS WITH CW OPERATION ABOVE 77K

Koji Shinohara, Mitsuo Yoshikawa,
Michiharu Ito and Ryuiti Ueda
Kansai Research Center
Fujitsu Laboratories Ltd., JAPAN

SUMMARY

Lead tin telluride diode lasers with emission wavelengths of 6 to 9 μm easily operate continuously at temperatures above 77K. These lasers have the $\text{Pb}_{1-y}\text{Sn}_y\text{Te}/\text{Pb}_{1-x}\text{Sn}_x\text{Te}/\text{Pb}_{1-y}\text{Sn}_y\text{Te}/\text{PbTe}$ (substrate), ($x > y$) double heterostructure.

To prepare this structure by LPE, the growth temperature must be below 600°C to suppress diffusion of the tin during the epitaxial growth.

When the heterojunctions are formed by the usual LPE method, the junction boundaries become irregular in the case for the lasers with wavelengths of over 10 μm at 77K.

In this paper, the mechanism by which the heterojunction boundaries become irregular is cleared and a new LPE method which prevents the irregularity is explained. The lasers prepared from the wafers grown by the new method have demonstrated CW operation at wavelengths longer than 10 μm above liquid nitrogen temperature.

INTRODUCTION

Lead tin telluride diode lasers are well known as excellent radiation sources in the wavelengths longer than 6 μm ⁽¹⁾. The diode lasers should be of double heterostructures to reduce the threshold current densities or to operate at temperatures above 77K⁽²⁾. Lasers with wavelengths from 6 to 10 μm operate easily above 77K and they are in practical use.

On the other hand, it is very difficult for the lasers to operate on CW with wavelengths over 10 μm above liquid nitrogen temperature. The reason for this is the difficulty of preparing flat heterojunctions for the lasers with wavelengths over 10 μm . The heterojunction boundaries become irregular if the junctions are made by the usual LPE method.

A new LPE method has been developed which overcomes the problem of irregular heterojunction boundaries.

By the new method, lasers capable of CW operation have been made at wavelengths longer than 10 μm at temperatures above 77K.

EXPERIMENTAL

Highly reliable lasers with mesa strata structure have been developed⁽³⁾⁽⁴⁾. The structure of the laser is shown in Figure 1. The heterojunctions have been prepared by the liquid phase epitaxial growth method using graphite sliding boat. To confine the photons and carriers in the active layer, the proportion of SnTe which is denoted as X in the active layer is higher than Y which is the proportion of two confining layers. The insulating films have been made by the anodic oxidation⁽⁴⁾.

The wavelength of the lasers is determined by the composition of the active region and the temperature. To make the wavelength longer than 10 μm at 77K, the value of X must be more than 0.2, while for wavelengths of 6 to 10 μm X is less than 0.2.

When the double heterojunctions are formed by the usual LPE technique, the shapes of the heterojunction boundaries vary obviously whether the composition of the active layer is over 0.2 or not.

Two photomicrographs of etched cross sections of double heterowafers are shown in Figure 2. Both are grown by the same LPE method. One is an example of the wafers for the 6 to 10 μm lasers, and the others is an example of the wafers for over 10 μm .

Heterojunctions contain many dislocations, which is due to the misfit of the lattice constant. The misfit dislocations between the substrate and the first layer spread in the direction of the crystal growth. This is due to the mutual self diffusion of Pb and Sn⁽⁵⁾. The growth temperature of the second and the third layers must be low enough to suppress mutual diffusion. Therefore, the active region is grown at a temperature of 600°C.

A photomicrograph for the long wavelength laser shows that the heterojunction boundaries between the first and second layers and between the second and third layers are irregular. When the heterojunction boundaries are irregular, the scattering loss that results becomes so serious that lasers made from wafers shown in Figure 2A do not operate even at 20K. But lasers with wavelengths from 6 to 10 μm , prepared from the wafer with flat heterojunctions shown in Figure 2B, give good CW operation above 77K.

To confirm the cause of irregularity in wafers for long wavelength lasers, another experiment was performed. The result is shown in Figure 3. The first layer of $Y = 0.13$ was grown all over the surface of the PbTe substrate. After removing the solution for the first layer, the solution for the second layer was kept on just half of the first epitaxial layer for 1 minute isothermally at 600°C. Then the solution was removed. The surface where the second solution was kept on is uneven, while the remaining half of the surface is smooth. The cross sectional view shows that at some places, the melting back occurs as was expected, and moreover, epitaxial growth occurs near the melted back area.

From the measurement of the lattice constant by the X-ray diffraction technique, the composition of the precipitated layer was determined to be $x = 0.25$.

From this experiment, the cause of irregularity is found to be due not only to the melting back but also the precipitation. Of course, the saturated solutions were used for the epitaxial growth, but even with these saturated solutions, both melting back and precipitation occurred.

The mechanism of the melting back and precipitation using saturated solutions can be explained by the phase diagram.

The ternary phase diagram of Pb, Sn and Te was investigated in detail by Harris, et al.⁽⁶⁾.

The temperature of the epitaxial growth of the second and third layers is below 600°C, so the region below the 600°C isothermal liquidus line must be considered. Note that the isocompositional solidus lines curve sharply where the composition is over $x = 0.2$.

First, the irregularity of the heterojunction boundary between the first and second epitaxial layers is considered.

The ternary phase diagram which is related to the first and second layers is shown in Figure 4. When the first layer is already grown and the solution for the second layer makes contact with the first layer, the first layer is in equilibrium with the solution whose composition is denoted by the point B where $Y = 0.13$ solidus intercepts the 600°C liquidus. So, the first layer is not in equilibrium with the solution at the point A where $x = 0.25$ solidus intercepts the 600°C liquidus, which is the composition of the second solution. The first composition and this second composition are so different that a disequilibrium emerges instantaneously when the second solution is held on the first epitaxial layer. To restore equilibrium between the solid and the liquid, the composition of the second solution changes from the point A toward the point B along the 600°C isothermal liquidus line.

In order to change the composition of this solution, the molecular interchange between the solid and the liquid of different composition is required.

By the local melting back of the first layer of $Y = 0.13$, the composition of the solution moves from the point A toward the point C which is the composition of the first layer. This compositional change is indicated by the vector a. On the other hand, by the precipitation of $x = 0.25$, the composition of the solution moves from the point A against the point D which is the composition of the second layer. This compositional change due to the precipitation is indicated by the vector b.

When the melting back and precipitation occur simultaneously, the motive force of the compositional change along the 600°C isothermal liquidus emerges. This is the reason why the interface of the first and second layers becomes irregular.

The mechanism of the irregularity between the second and the third layers is just the reverse mechanism of the aforementioned irregularity at the interface of the first and second layer; that is, the composition of the third solution moves toward the composition of the second solution along the 600°C liquidus line.

All this explains why the heterojunctions become irregular when double heterowafers are produced by the usual LPE method.

However, a way has been found to produce flat heterojunction boundaries on double heterowafers for the 6 to 10 μm lasers.

The reason this can be done can be understood from the phase diagram. The isocompositional lines are straight below $x = 0.20$. So, the compositional difference of the solutions between the first layer and the second layer is not so large, the result being that serious disequilibrium does not occur. Of course, the solution composition for the second layer moves toward that of the first solution, but the motive force is so small that the boundaries do not become irregular.

From the aforementioned analysis, it is seen that to obtain flat heterojunction boundaries for the long wavelength laser, it is necessary to produce precipitation only, while suppressing the melting back. So, supercooled solutions for the growth of the second and third layers were used in order to allow no time for compositional change.

The temperature profile for the usual LPE growth method is shown in Figure 5. The first layer is grown slowly to assure the quality of the epitaxial layer, then the second and third layers are grown continuously at the cooling rate of 1°C per minute.

The modified temperature profile in order to achieve the supercooled condition is shown in Figure 6. The beginning half of the first layer is grown under the usual conditions, but in the middle point of the first layer growth, the cooling rate is changed to as much as 3°C per minute. While the last half of the first layer is growing, the solutions for the second and third layers are sufficiently supercooled prior to making contact with the substrate. Then the second and the third layers are grown continuously.

The cooling rate of 3°C per minute is sufficient to keep the solutions supercooled.

A photomicrograph of the cross section grown by the new LPE method is shown in Figure 7. Notice how flat the heterojunction boundaries are.

Lasers prepared from the wafers like this easily operate on CW with wavelengths longer than 10 μm above 77K.

In Figure 8, one example of the lasing spectrum is shown demonstrating CW operation on an 11.45 μm wavelength with single mode. The temperature of the heat sink is 77K.

CONCLUDING REMARKS

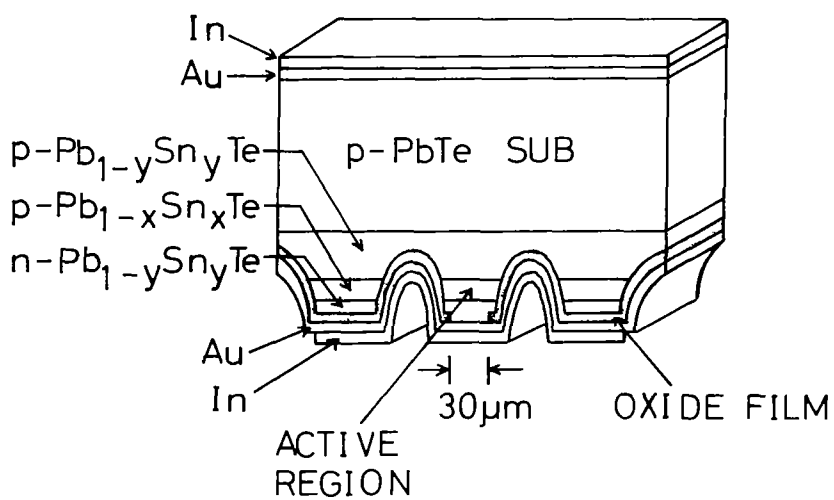
The mechanism by which heterojunction boundaries become irregular was shown and a new method was explained of liquid phase epitaxial growth which prevents this irregularity.

Lasers prepared from the wafers grown by this new LPE method have demonstrated CW operation at wavelengths longer than 10 μm above 77K.

We assume that lasers with emission wavelengths longer than 16 μm can also be operated continuously above 77K if necessary.

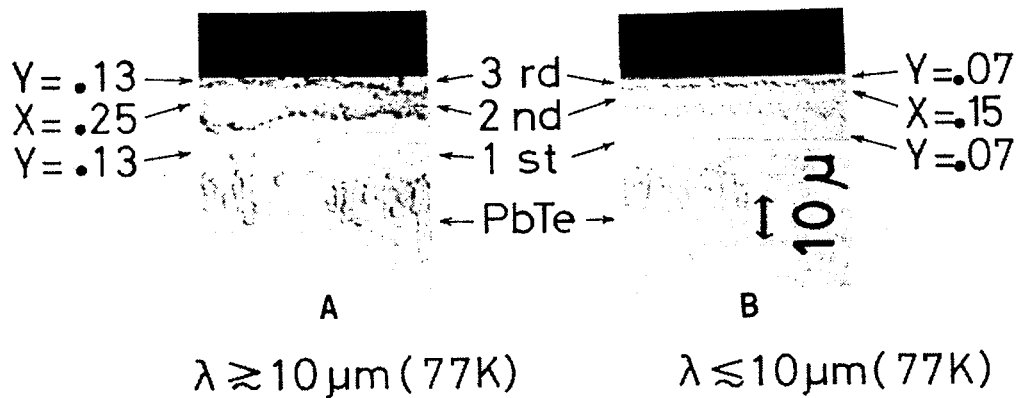
REFERENCES

1. Dimmoch, J. O.; Melngailis, I.; and Strauss, A. J.: Phys. Rev. 16, 1193 (1966)
2. Groves, S. H.; Nill, K. W.; and Strauss, A. J.: Appl. Phys. Lett. 25, 331 (1974)
3. Yoshikawa, Mitsuo; Shinohara, Koji; and Ueda, Ryuiti: Appl. Phys. Lett. 25, 699 (1977)
4. Yoshikawa, Mitsuo; Koseto, Masaru; and Ueda, Ryuiti: CLEOS '78, San Diego
5. Yoshikawa, Mitsuo; Ito, Michiharu; Shinohara, Koji; and Ueda, Ryuiti: J. Crystal Growth 47, 230 (1979)
6. Harris, J. S.; Longo, J. T.; Gertner, E. R.; and Clarke, J. E.: J. Crystal Growth 28, 334 (1975)



$X \geq 0.2$ for $\lambda \geq 10\mu\text{m}$
 $X \leq 0.2$ for $6 \leq \lambda \leq 10\mu\text{m}$

Figure 1.- Schematic structure of the laser diode.



A is an example of the wafers for over $10 \mu\text{m}$ at 77K

B is an example of the wafers for 6 to $10 \mu\text{m}$

Figure 2.- Etched cross sections of double heterowafers.

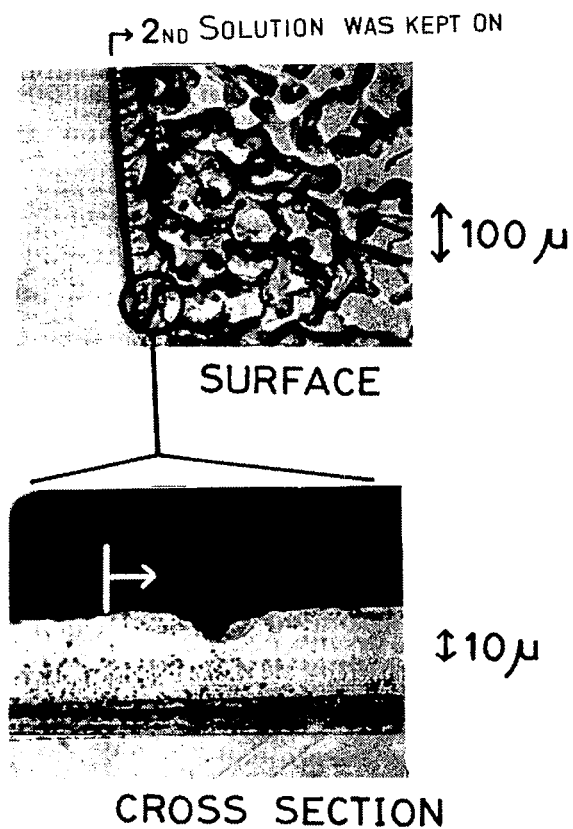


Figure 3.- Heterointerface morphology.

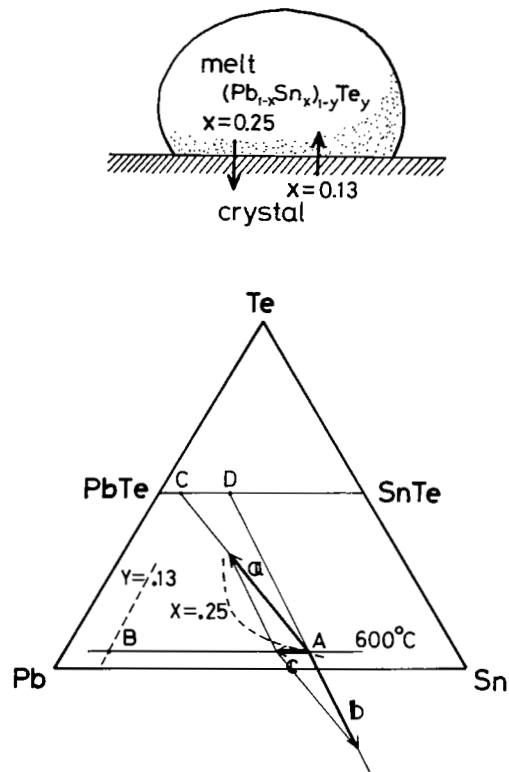


Figure 4.- Ternary phase diagram of Pb-Sn-Te explaining the heterojunction boundary's irregularity.

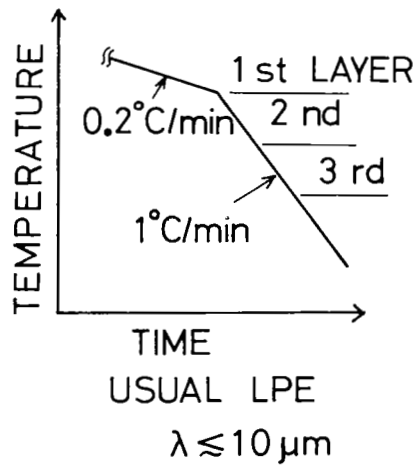


Figure 5.- The temperature profile of usual LPE method.

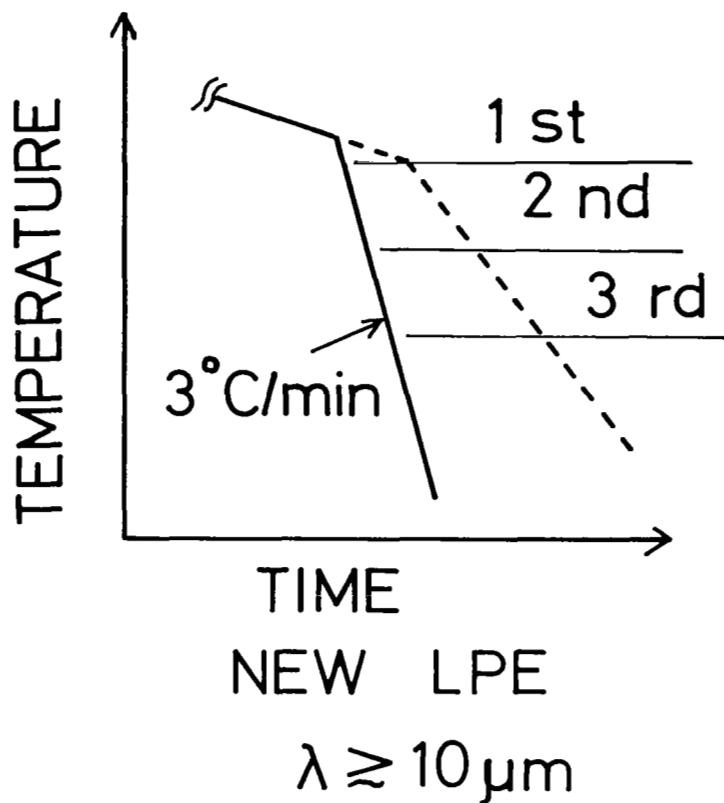


Figure 6.- The temperature profile of the new LPE method.

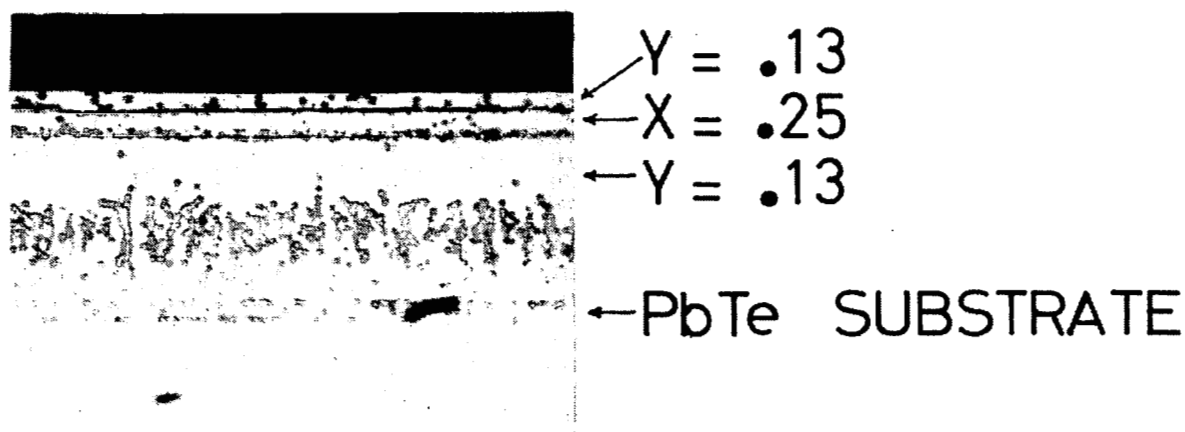


Figure 7.- Etched cross section of the double heterostructure for the long wavelength lasers grown by the new LPE method.

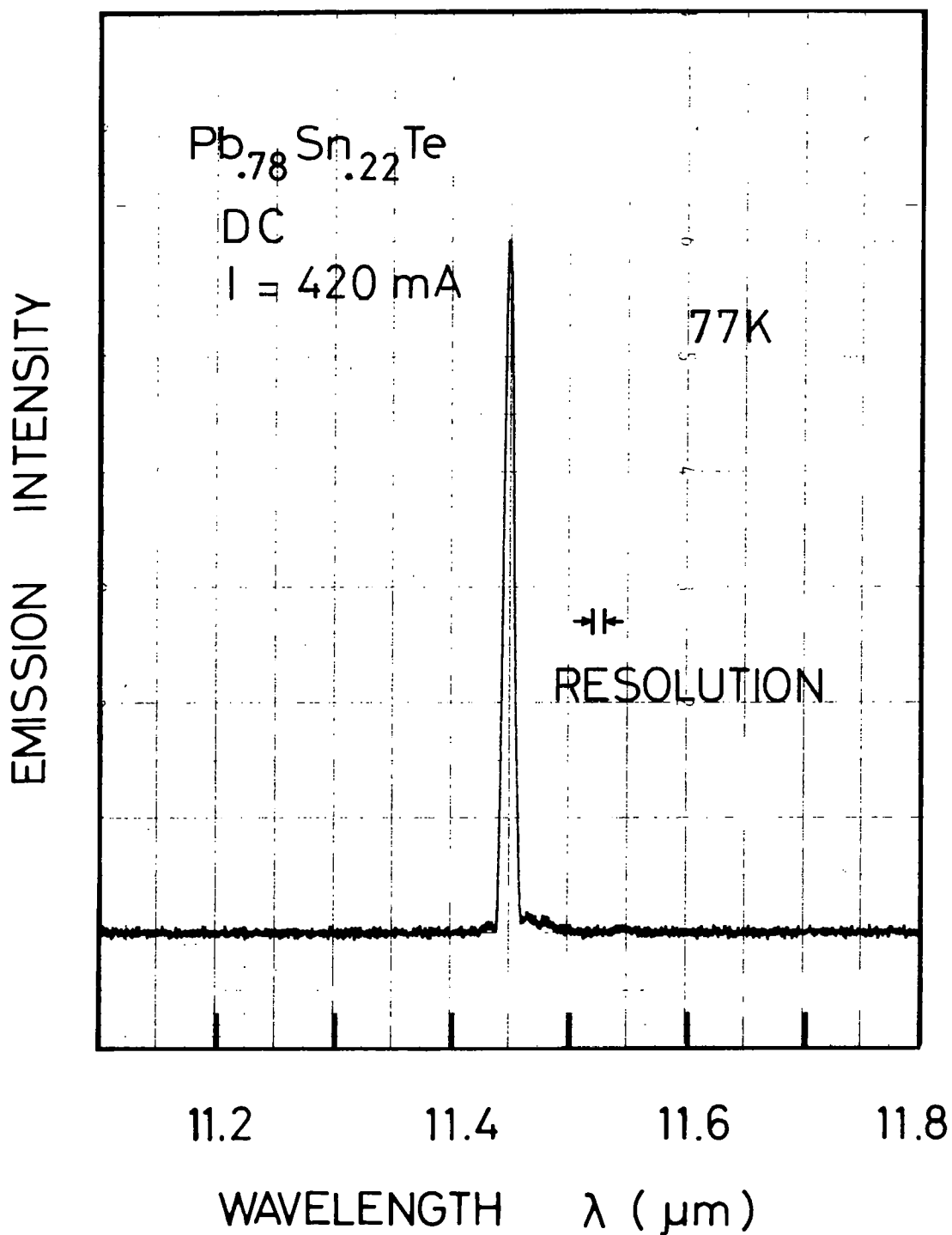


Figure 8.- One example of the lasing spectrum.

WIDELY TUNABLE (PbSn)Te LASERS USING ETCHED CAVITIES FOR MASS PRODUCTION[†]

Matthew D. Miller
The Perkin-Elmer Corporation

SUMMARY

Lead salt diode lasers are being used increasingly as tunable sources of monochromatic infrared radiation in a variety of spectroscopic systems. These devices are particularly useful, both in the laboratory and in the field, because of their high spectral brightness (compared to thermal sources) and wide spectral coverage (compared to line-tunable gas lasers). While the primary commercial application of these lasers has been for ultra-high resolution (10^{-4} cm^{-1}) laboratory spectroscopy, there are numerous systems applications, including laser absorption pollution monitors and laser heterodyne radiometers, for which diode lasers have great potential utility. There are, however, several problem areas related to the wider use of these components. Among these are total tuning range, mode control, and high fabrication cost. A fabrication technique which specifically addresses the problems of tuning range and cost, and which also has potential application for mode control, is reported here.

PROCEDURE

Crystal Growth

(Pb_{0.81}Sn_{0.19})Te homojunctions were grown by multiple source molecular beam epitaxy (MBE). Independently controlled sources of PbTe and SnTe permitted precise composition control, while Bi₂Te₃ and Tl were used as extrinsic n and p dopants, respectively. A fifth Knudsen source was included to provide excess Te, if necessary, to control stoichiometry. The substrate was BaF₂, first used by Holloway (ref. 1). Unlike earlier work which used the (111) cleavage plane, the substrates for this work were chemically polished (100) oriented material, to allow for the preference of the lead salts to grow in the (100) orientation.

Crystal growth took place at temperatures from 400 to 420°C and growth rates of from 2 to 4 μm per hour. The laser's vertical structure was a five-layer homojunction utilizing carrier concentration gradients for both optical and electrical confinement. Structures similar to this have been reported by Lo (ref. 2) using bulk diffused material and by Walpole (ref. 3) using MBE-grown material on a bulk substrate. A typical growth sequence consisted of growing an n-type contact layer 0.5 μm thick and doped to 10^{19} cm^{-3} , an

[†]Research partially supported by NASA contract NAS1-14996

n-type buffer layer 2 μm thick and doped to $3 \times 10^{18} \text{ cm}^{-3}$, an undoped active layer 1.5 μm thick which was typically 1×10^{18} p-type, 2 μm of 3×10^{18} p-type buffer, and, finally, 0.5 μm of 1×10^{19} p-type contact layer. Following growth, the entire multilayer film, which was typically 8mm long, 10mm wide, and 9 μm thick, was metallized and soldered to a copper heat sink, forming electrical contact to the p-contact layer. The BaF_2 substrate was then slowly dissolved off using warm flowing de-ionized water.

Device Fabrication

Standard integrated circuit processes were used to mass-produce mesa diode lasers. The process steps from crystal growth to wire bonding are summarized in figure 1. Following substrate removal, rectangular Fabry-Perot cavities were etched, either wet-chemical or by ion milling. Next, a BaF_2 insulating layer was deposited and a contact window opened to permit electrical contact to the n-type contact layer without creating an electrical short to the copper heat sink. Metal was then deposited over the entire structure and subsequently removed from one face of the laser. The device was completed by wire-bonding to the top contact metal, next to the device, in order to avoid damage due to mechanical stress. Figure 2 shows the finished mesa diode in detail.

A typical structure is 75 μm wide, 250 μm long, and 9 μm high. An array of these diodes prior to final wire bonding is seen in figure 3. The diodes are spaced 500 μm apart in x and y. The dark squares are the photoresist which has been used to pattern the top contact metal, and the smaller rectangles within the squares are the lasers, where one face of the laser is covered by metal and one face extends beyond the metal, permitting radiation to escape.

LASER PERFORMANCE

Diodes fabricated in this way have been operated cw at heat sink temperatures ranging from 12K to 60K, and they operated in 5 μs pulses up to 95K, at which temperature the pulsed threshold current was $10,000 \text{ A/cm}^2$. Low temperature thresholds were between 500 A/cm^2 and 1000 A/cm^2 , both pulsed and cw. These values for threshold currents compare favorably with those obtained for devices made by more conventional one-at-a-time processes based on the growth of high quality bulk single crystal material.

The emission wavelength of lead salt lasers is varied by changing the temperature at the junction. This tuning is accomplished by varying the heat sink temperature and the current in the laser, together or separately. A laser can operate in either a single mode or multimode, depending upon small changes in the operating point. Figure 4 illustrates both the current tuning and the mode structure of a diode operated at a heat sink temperature of 20K. Since the average emission wavelength is determined primarily by the temperature dependence of the band gap of the semiconductor, one can predict the average wavelength of emission given the I-V characteristics of the laser, the measured thermal resistance, and the known temperature dependence of the band gap.

Figure 5 summarizes both the temperature tuning and the current tuning of a laser which operated in the pulsed mode up to 95K. The cw emission wavelength is taken to be the center frequency for cases of multimode emission. This wavelength is plotted as a function of current and heat sink temperature. The solid lines were calculated from the known electrical and thermal properties of the diode.

Because of difficulties in the thermal package, this diode exhibited quite a high temperature rise at the junction; thus, the cw emission of 10 μm at 1.35A and a heat sink temperature of 15K corresponded to a junction temperature of 95K, the highest temperature for pulsed operation. Significantly, we have obtained, in a single diode, cw emission from 10 μm to 14 μm , a tuning range comparable to the widest reported in the literature.

CONCLUSION

Useful laser performance has been demonstrated in devices fabricated using thin-film and integrated circuit technologies geared toward mass production. The MBE crystal growth technique permits independent and precise control of composition and carrier concentration. The lasers described in this paper were grown on a BaF_2 substrate and exhibited an extremely wide tuning range, despite a 1.4% lattice mismatch. This wide tuning range is probably attributable to excellent optical and electrical confinement, accomplished with the thin active region and sharp concentration gradients achievable with MBE.

On the other hand, these devices did not show good performance in terms of single mode or multimode output power. There are two likely causes for the low output power (typically 10 μW). There is probably non-optimal output coupling due to imperfect and misaligned etched end mirrors. Refinements of the ion-milling process have led to some improvements in this area. More significantly, there may be severe quantum efficiency reduction due to the lattice mismatch. This effect has been demonstrated by Kasemset for $(\text{PbSn})\text{Te}$ heterojunctions (ref. 4).

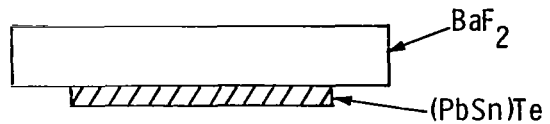
Figure 6 presents the range of lattice constants for some potential substrate materials. The only possibilities for lattice-matching to $(\text{PbSn})\text{Te}$ are bulk material or the mixed alkali-halide, which is a very poor thermal match and not a very good substrate. On the other hand, it is possible to lattice-match $(\text{PbSn})\text{Se}$ to $(\text{BaSr})\text{F}_2$.

In addition to potential cost reduction, the thin-film processing techniques provide the possibility for integrated arrays of diodes differing from each other in a controlled fashion. This would greatly aid in the problem of spectral coverage. The photolithographic cavity definition allows the use of shaped cavities for mode control, and the multi-layer thin film crystal growth permits great flexibility in vertical structure. Thus, standard devices may be obtained with improved properties in terms of both cost and performance.

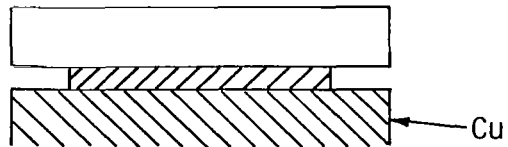
REFERENCES

1. Holloway, H. et al.: Injection Luminescence and Laser Action in Epitaxial PbTe Diodes. Appl. Phys. Lett., vol. 21, no. 5, 1972.
2. Lo, W.: Homojunction Lead-Tin-Telluride Diode Lasers with Increased Frequency Tuning Range. IEEE Journal of Quantum Electronics, QE-13, p. 591, 1977.
3. Walpole, J.N. et al.: MBE Homostructure PbTe Diode Lasers with cw Operation up to 100K. Device Research Conference, Ithaca, NY, June 1977.
4. Kasemset, D. and Fonstad, C.: Minority Carrier Lifetimes and Lasing Thresholds of PbSnTe Heterostructure Lasers. IEEE Transactions on Electron Devices, ED-26, p. 1841, 1979.

1. FILM GROWN BY MULTIPLE SOURCE MBE ON BaF_2 SUBSTRATE



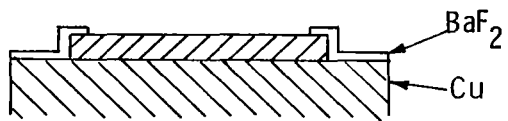
2. SOLDERED TO COPPER HEAT SINK



3. BaF_2 DISSOLVED WITH WARM WATER, FABRY-PEROT CAVITIES ETCHED



4. BaF_2 INSULATOR DEPOSITED, CONTACT WINDOWS OPENED



5. TOP METALIZATION DEPOSITED, REMOVED FROM EMITTING FACE, CONTACT WIRE BONDED AWAY FROM ACTIVE LASER

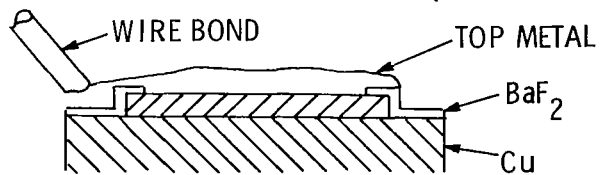


Figure 1.- Processing steps in laser fabrication, from crystal growth to final wire-bonding.

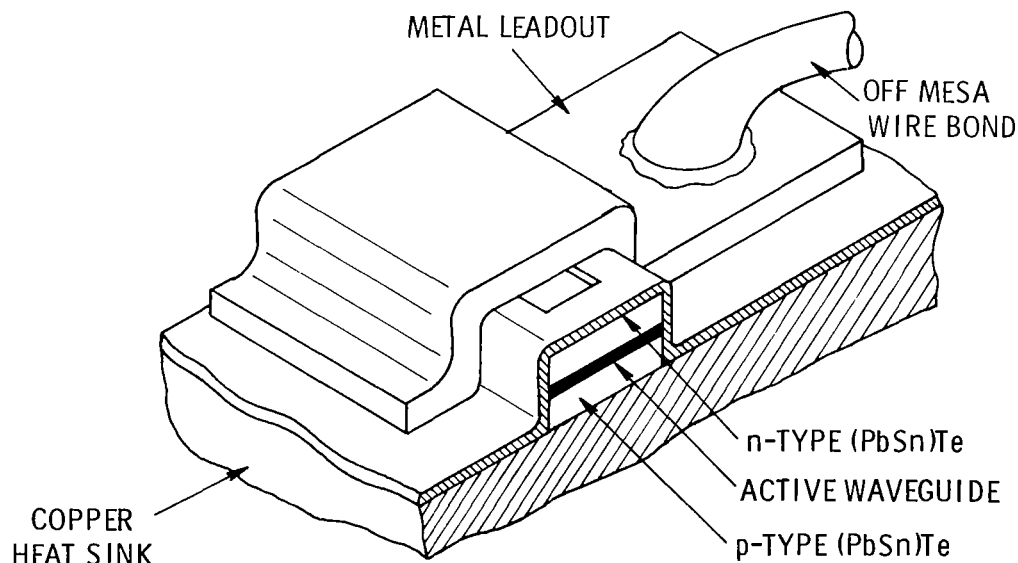


Figure 2.- Detailed illustration of etched mesa diode laser.



Figure 3.- Scanning electron micrograph of an array of lasers prior to wire-bonding.

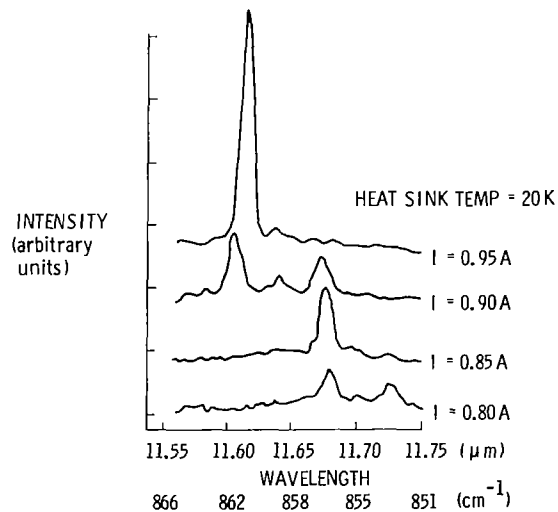


Figure 4.- Output spectrum illustrating complex mode structure and rapid wavelength tuning with current.

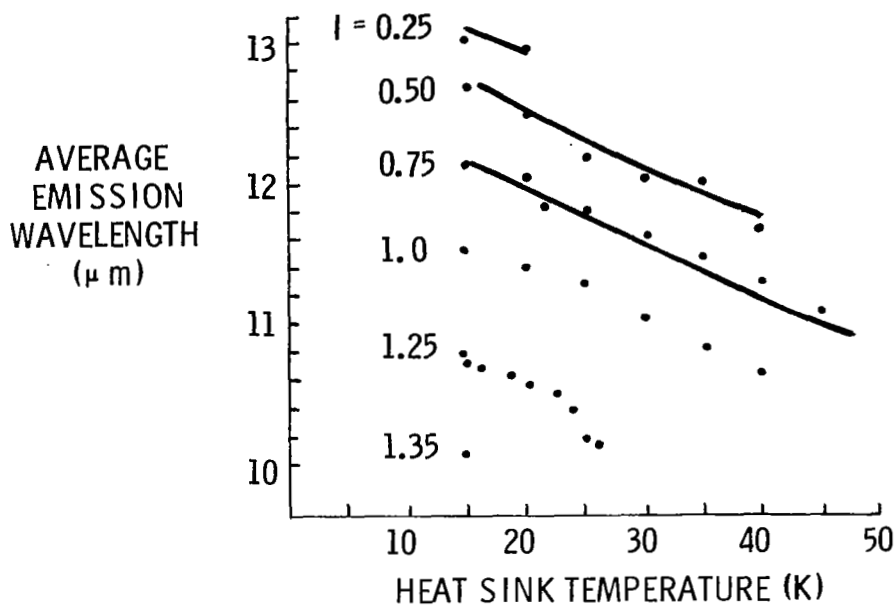


Figure 5.- Range of accessible wavelengths for diode operating cw. Wavelength tuned with temperature and current from 14 μm to 10 μm . Solid lines are theoretical tuning curves derived from temperature dependence of the band gap and known electrical and thermal properties of the diode.

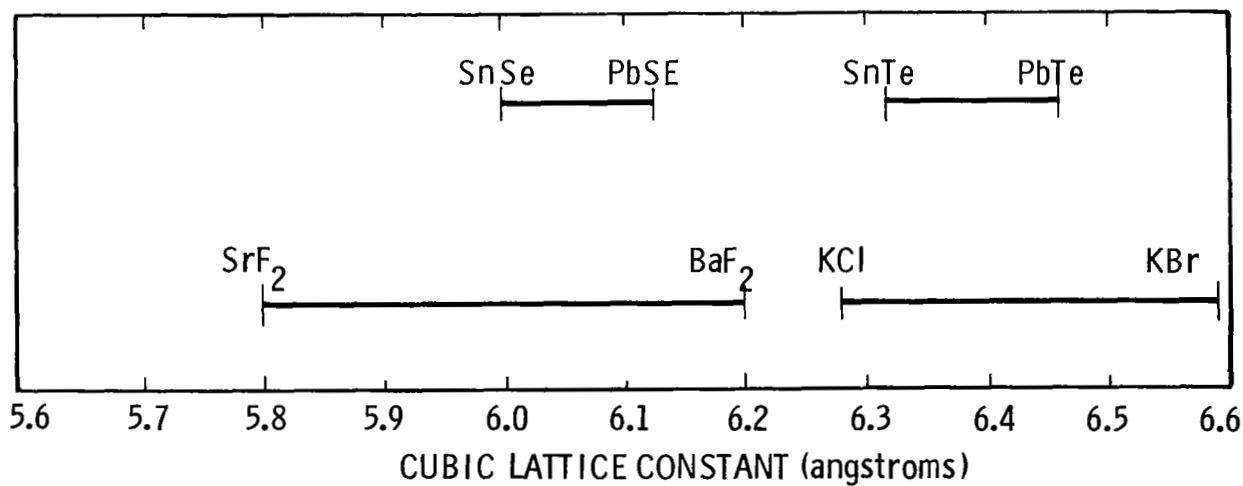


Figure 6.- Range of cubic lattice constants for potential substrate materials for epitaxial growth of lead salt laser material.

RELIABILITY IMPROVEMENTS IN TUNABLE $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ DIODE LASERS

OPERATING IN THE 8.5-30 MICROMETER SPECTRAL REGION*

Kurt J. Linden, Jack F. Butler, Kenneth W. Nill and Robert E. Reeder
Laser Analytics, Inc.

SUMMARY

This paper describes recent developments in the technology of Pb-salt diode lasers which have led to significant improvements in reliability and lifetime, and to improved operation at very long wavelengths. A combination of packaging and contacting-metallurgy improvements has led to diode lasers that are stable both in terms of temperature cycling and shelf-storage time. Lasers cycled over 500 times between 77K and 300K have exhibited no measurable changes in either electrical contact resistance or threshold current. Utilizing a newly developed metallurgical contacting process, both lasers and experimental n-type and p-type bulk materials have been shown to have electrical contact resistance values that are stable for shelf storage periods well in excess of one year. This paper summarizes some of the previously encountered problems and reviews several experiments which have led to devices with improved performance stability. Such stable device configurations have been achieved for material compositions yielding lasers which operate continuously at wavelengths as long as 30.3 μm .

INTRODUCTION

A number of different failure mechanisms have been observed in Pb-salt diode lasers. Some of these, such as failures due to power supply transients, accidental back-biasing or excessive forward-bias currents are operator or equipment related and can be avoided by the user if manufacturer's instructions are followed. Other failures observed appear to have been related to either temperature cycling (ref. 1) or shelf storage of the lasers (ref. 2). It is such failures and the successful efforts to overcome them that are described in this paper. In contrast to observations made on some of the III-V compound diode lasers, laser degradation under operating conditions has not been a problem with the Pb-salt lasers.

Stability in the operating characteristics of diode lasers requires stability in all parameters affecting these characteristics. Such parameters include the device mechanical characteristics (i.e., temperature-dependent

*The work reported here has, in part, been supported by the NASA Langley Research Center and the Los Alamos Scientific Laboratory. Some of the results on thermal resistance stability were obtained on a program supported by the U.S. Army Harry Diamond Laboratories.

stress effects which are usually package related), electrical characteristics (i.e., the stability of the electrical contact resistance) and thermal characteristics (i.e., the stability of the device thermal resistance). This paper describes the successful stabilization of the first two parameters. The problem of thermal resistance stability has been studied as part of a special research effort under Army contract DAAK21-79-C-0041, portions of which are reported here.

Following a description of present laser packaging techniques, the thermal cycling stability of some typical $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ diode lasers is described in the section entitled "Thermal Cycling Stability". The section entitled "Previously Observed Contact Resistance Degradation" discusses some typical contact resistance degradation problems observed in both p-type and n-type $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ materials as well as in p-n junction lasers. The section entitled "Diagnostic Measurements and Stability Results on Lasers Fabricated by Newly Developed Techniques" describes experiments aimed at investigating sources of such degradation and presents results on the stability of bulk n-type and p-type devices and of lasers fabricated by newly developed techniques. Some special results obtained for long-wavelength lasers of $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ are presented in the section entitled "Long Wavelength Laser Performance and Stability". Finally, an effort aimed at separating the effects of electrical and thermal resistance has led to some interesting observations, as described in the section entitled "Thermal Resistance Stability Measurements".

SYMBOLS

R_f	forward bias resistance of diode lasers (ohms)
I_{th}	laser threshold current (mA)
ΔP	relative quantum efficiency of laser (no units)
P_{max}	relative maximum output power of laser (no units)
RA	resistance - active area product of diode laser
λ	free space wavelength (μm)
W	power (mW or μW as indicated)
CW	continuous wave
I	laser current (mA)
n	refractive index of laser material
L	length of optical cavity of the diode laser (μm)
m	order of longitudinal mode (an integer)
T	temperature (K)

THERMAL CYCLING STABILITY

The temperature extremes to which Pb-salt lasers are subjected are among the most severe in the semiconductor industry. Ranging from liquid helium (4.2K) to room temperature (300K), such temperature extremes require careful choice of crystal mounting techniques, package materials and physical configurations. The choice of materials is determined by such factors as thermal conductivity, thermal expansion coefficient and electrical properties at cryogenic temperatures.

The presently used package for Pb-salt diode lasers is shown in the scanning electron microscope (SEM) photograph of figure 1. This package provides easy access to the laser face and utilizes a pressure-free design in which the laser crystal is welded between an indium-plated stud and a flexible top contact strap.

In an experiment to investigate the temperature cycling stability of this package, 3 $\text{Pb}_{1-x}\text{Sn}^x\text{Se}$ diode lasers were cycled between room temperature and 77K a total of 500^x times. Each cycle consisted of a 10 minute dip in liquid nitrogen followed by a 20 minute blower-driven warm-up, with the lasers mounted in an evacuated "dip-stick". The dip-stick was designed to facilitate electrical testing by dipping into a liquid helium tank. Both the electrical contact resistance and threshold current were periodically measured. The results are summarized in table I. The variations in threshold current are commonly observed in Pb-salt diode lasers and are believed due to changes in surface leakage current. The contact resistance is essentially the same over the entire test period, since a variation of $\pm .0005$ ohms is well within the expected experimental error. It is thus apparent that 500 temperature cycles have had no significant effect on the laser contact resistance. Subsequent room temperature storage lifetime data on these lasers are presented in the section entitled "Diagnostic Measurements and Stability Results on Lasers Fabricated by Newly Developed Techniques".

PREVIOUSLY OBSERVED CONTACT RESISTANCE DEGRADATION

The electrical contact resistance to diode lasers is a major factor in laser heating and therefore current tuning rate during operation. A common source of shelf storage degradation of Pb-salt diode lasers has been a progressing increase in this parameter. This phenomenon was even found to occur on the bulk n-type and p-type $\text{Pb}_{1-x}\text{Sn}^x\text{Se}$ samples. The contact resistance to n-type and p-type samples of $\text{Pb}_{0.926}\text{Sn}_{0.074}^x\text{Se}$, measured over a period of several months, is shown in figure 2. The sample areas were approximately $8 \times 10^{-4} \text{ cm}^2$ and the storage conditions were at room temperature, under ambient conditions.

The progressive contact resistance degradation illustrated here was somewhat erratic in that not all lasers degraded this rapidly. Lasers can, in general, operate under CW conditions with contact resistance values as high as approximately 0.1 ohms. Thus lasers with initial resistance values of 0.005 ohms may continue operating for long periods before the resistance increases to the point where CW laser action ceases. During this time, however, the progressive resistance increases would cause continual changes in the laser tuning rates and operating frequencies at given temperature-current combinations.

The diagnostic measurements carried out to investigate the sources of this contact resistance degradation and the results of these measurements are discussed in the following section.

DIAGNOSTIC MEASUREMENTS AND STABILITY RESULTS ON LASERS

FABRICATED BY NEWLY DEVELOPED TECHNIQUES

The existence of progressive increases in electrical contact resistance of both bulk and p-n junction devices suggested metallurgical interactions either between the contact metal layers or between the semiconductor and the metal layers as possible sources of this problem. These possibilities were investigated by use of Auger electron spectroscopy (AES) of the semiconductor and metal interface surfaces. Several $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ wafers were analyzed by removing the metallization layer by a tape-pulling technique and probing the semiconductor surfaces. The results are summarized in Table II. These results suggest that failing devices were characterized by the presence of In on the p-type semiconductor surface extending at most 150\AA below the crystal surface (such a distance being the resolution limit of these AES measurements). Since In is a critical part of the bond used in mounting the laser crystals into the laser packages, the elimination of this material was not considered. To prevent the apparent migration of In through the Au and, in some cases, barrier metal layer, the use of new metallic barrier layers was investigated. The proper use of such layers can be expected to prevent In migration to the crystal surface. A process utilizing multiple metallization layers deposited by electron beam evaporation techniques was developed, and a number of experimental bulk n-type and p-type devices were fabricated and evaluated for shelf storage stability. These devices have exhibited excellent contact resistance stability to date. As of 2/20/80, the p-type samples have an accumulated shelf storage time of 17 months, while the n-type samples have been shelf-stored for almost 15 months. The retest data for these junctionless devices are summarized in figure 3. As the device areas were approximately $8 \times 10^{-4} \text{ cm}^2$, the RA products of the n-type samples are approximately $4 \times 10^{-7} \text{ ohm cm}^2$ per side while those of the p-side are approximately $1.6 \times 10^{-6} \text{ ohm cm}^2$. The n-contact values are among the lowest ever reported to any semiconductor material.

The stability of a number of CID-type (ref. 3) $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ lasers fabricated by the improved metallization techniques was tracked for a period of over a year. Some of these results are shown in figure 4 which tabulates the

contact resistance, threshold current and relative quantum efficiency as a function of time for a group of 4 $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ lasers. A total time period of 15 months is covered by these data (the latest points corresponding to measurements made on 2/20/80). The relative quantum efficiency parameter was obtained by measuring the relative rate at which the output power increased with laser current just above threshold. In absolute terms, the external quantum efficiencies corresponding to these values are of the order of 0.5% per laser end. The comparative data are presented here to show that the output power from these lasers was relatively stable over the 15 month shelf storage time period. All data points presented in figure 4 were obtained at liquid helium temperature. The emission spectra for one of these lasers (8292-2) as measured under identical test conditions at time periods approximately 1 year apart are shown in figure 5. While there appears to be a spectral shift towards higher energy by about 6 cm^{-1} , the relative longitudinal mode distribution appears quite stable, and the total output power somewhat higher. Some of this discrepancy may be due to a change of power meters during this time interval.

A second group of lasers stored and retested over a 15 month period are the 3 devices which comprised the group described in the section entitled "Thermal Cycling Stability". After being temperature recycled 500 times the devices were periodically retested for contact resistance, threshold current, relative quantum efficiency (ΔP) and relative maximum output power (P_m). As in the case of relative quantum efficiency, the relative maximum output power corresponded to the current through a Ge:Cu photodetector mounted adjacent to the laser and is shown for comparative purposes only. The actual maximum output power for these lasers was of the order of a mW per end. The value of these test parameters at various retest intervals are shown in table III. All parameters are seen to be relatively stable and show no tendency for degradation over this time period. Unit 8324-10 apparently had an initially high threshold current (which could have been an instrument error) but thereafter appears to have settled down. The period in time at which these devices were temperature cycled is indicated in the table.

A third group of lasers consisted of 2 units which were fabricated from $\text{Pb}_{0.928}\text{Sn}_{0.072}\text{Se}$ (corresponding to $16\text{ }\mu\text{m}$ emission at 15K). These units were periodically retested over an 11 month shelf storage period and then delivered under NASA-Langley contract NAS1-15190. The results of these measurements, shown in table IV, indicate excellent stability.

The metallization process used in fabricating the devices described here was shown to be capable of withstanding air-ambient bake temperatures of 65°C for periods of 63 hours. In an experiment involving 9 lasers from several different $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ wafers, 8 of them (89%) survived this bake. Of these survivors, 6 (or 75%) remained stable thereafter. While it is not standard practice to subject Pb-salt lasers to such high temperatures, the experiment demonstrated the relative thermal stability of the presently used metallization process.

LONG WAVELENGTH LASER PERFORMANCE AND STABILITY

Long wavelength lasers ($\lambda > 25 \mu\text{m}$) have been fabricated from $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ material with compositions on both side of the zero-gap crossover point (ref. 4). The temperature dependence of the bandgap (hence, emission frequency) on opposite sides of the crossover at $x \approx 0.14$ differ markedly. For $x < 0.14$ the band gap (emission frequency) increases with increasing temperature while for $x > 0.14$ the band gap decreases with increasing temperature. We report the operational results obtained from some $\text{Pb}_{0.90}\text{Sn}_{0.10}\text{Se}$ lasers in the 29-30 μm spectral region. This composition region has the advantage of yielding the most power at the long wavelength region of operation, since this occurs at the lowest temperatures. The emission spectrum of a $\text{Pb}_{0.90}\text{Sn}_{0.10}\text{Se}$ laser operating CW at 10K and 2000 mA is shown in figure 6. With a total measured output power of 141 μW , the mode which occurs at 338.5 cm^{-1} (29.5 μm) has a CW power of approximately 30 μW . When operated just above threshold, the laser emits at 330 cm^{-1} (30.3 μm) in a single mode. This spectrum is shown in figure 7.

The stability of four such lasers was monitored over a 10 1/2 month time period (as of 2/22/80). A plot of contact resistance, threshold current and relative quantum efficiency is shown in figure 8. All operating parameters are reasonably stable over this time period, although two of the lasers do exhibit some increase in contact resistance. As in the measurements described previously, some of the differences in the relative quantum efficiency are due to the fact that these measurements were made in different sample test fixtures with different detector sensitivity values.

THERMAL RESISTANCE STABILITY MEASUREMENTS

In an effort to determine the thermal resistance stability of Pb-salt diode lasers, several units known to have stable electrical contact resistance values were checked for single mode tuning rate. This parameter was periodically remeasured. The results indicated that while some devices are stable, some can also show a slight tendency towards increasing thermal resistance with time.

The single mode tuning rate $\frac{\partial(\frac{1}{\lambda})}{\partial I}$ is given by

$$\frac{\partial(\frac{1}{\lambda})}{\partial I} = \frac{m}{2nL} \left(\frac{1}{n} \frac{\partial n}{\partial T} + \frac{1}{L} \frac{\partial L}{\partial T} \right) \frac{\partial T}{\partial I} \quad (1)$$

where m is the order of the longitudinal laser mode, typically 550 for a 1/2 mm cavity length, L, at 10 μm . The refractive index, n, is temperature dependent and the $\frac{1}{n} \frac{\partial n}{\partial T}$ term is larger than the thermal expansion term $\frac{1}{L} \frac{\partial L}{\partial T}$ (which is known to be approximately $20 \times 10^{-6} \text{ K}^{-1}$ for PbSe). The

$\frac{\partial T}{\partial I}$ term is proportional to the thermal resistance of the laser. Thus, the single mode tuning rate provides a qualitative measure of the thermal resistance. Preliminary measurements indicate that small thermal resistance increases can occur on devices that are electrically stable. This technique provides a useful tool for analyzing the thermal stability of various laser mounting techniques.

An experiment was performed to measure the temperature differential across a laser crystal under operating conditions. By use of a precalibrated miniature Si diode sensor which was mounted directly to the top contact strap of a laser, it was possible to measure the temperature at that point, while the temperature on the other side (in contact with the large heat sink of the mounting stud) was measured with a second calibrated sensor. It was found that temperature differentials of 2-5K were typical at the full laser operating current values of 2 amps. To the best of our knowledge, this represents the first reported measurement of this parameter.

CONCLUDING REMARKS

Package design modifications and contact technology developments have led to significant improvements in the reliability and stability of $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ diode lasers operating in the 8-30 μm spectral region. Representative lasers subjected to over 500 temperature cycles between 77K and 300K have exhibited no significant changes in their operating parameters. Lasers stored under ambient temperature conditions have exhibited excellent stability in electrical contact resistance, threshold current and relative output power over test periods of 15 months. These improvements have been successful even for long wavelength material compositions with the demonstration of laser performance stability for periods in excess of 10 months to date. Such lasers have emitted up to 30 μW of CW power at 29.5 μm . Recently fabricated lasers have been shown capable of withstanding 65°C bake temperatures with high survival and good subsequent stability yields.

Indirect measurements of thermal resistance have indicated that it is, in some cases, possible to measure small increases in thermal resistance on devices whose electrical contact resistance is stable. Such findings indicate the diverse nature of the problems associated with diode laser failure mechanisms.

REFERENCES

1. M. Yoshikawa, K. Shinohara and R. Ueda, Appl. Phys. Lett. 31, 699 (1977).
2. H. Preier, Appl. Phys. 20, 189 (1979).
3. K. J. Linden, K. W. Nill and J. F. Butler, IEEE Jour. Quant. Elec., QE-13, 720 (1977).
4. T. C. Harman, A. R. Calawa, I. Melngailis and J. O. Dimmock, Appl. Phys. Lett. 14, 333 (1969).

TABLE I.-STABILITY DATA ON LASERS CYCLED 500 TIMES BETWEEN 300K AND 77K.

Laser No.	Original Data		After 135 temp cycles		After 500 temp cycles	
	R_f (ohms)	I_{th} (mA)	R_f (ohms)	I_{th} (mA)	R_f (ohms)	I_{th} (mA)
8324-11	.005	1410	.006	1360	.005	1320
8324-12	.006	1270	.005	1200	.005	1200
8324-10* (control)	.005	> 2000	.005	1380	.005	1350

*Laser 8324-10 was a control sample from the same lot. This laser was not temperature cycled.

TABLE II.-SUMMARY OF AES FINDINGS ON $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ LASER WAFERS.

Wafer No.	Description	AES Findings
D700 p-side	P-diffused, striped wafer of $\text{Pb}_{.93}\text{Sn}_{.07}\text{Se}$. Lasers exhibited contact resistance increases after a few months.	<u>Surface:</u> Trace of Au (~.1%). Strong In signal (~few %). More In in stripe than on insulator. <u>Below Surface:</u> 50 Å of material removed gave same results as above. 150 Å of material removed, and the In signal was gone.
D1002 p-side n-side	N-diffused, non-striped wafer of $\text{Pb}_{.989}\text{Sn}_{.011}\text{Se}$. Lasers degraded in a few days.	<u>Surface:</u> Au approximately 2% In approximately 2% <u>Surface:</u> Approximately 5% In
D448 p-side	N-diffused, striped wafer of $\text{Pb}_{.935}\text{Sn}_{.065}\text{Se}$. Lasers were stable for over one year.	<u>Surface:</u> Approximately 5% Au, no trace of In found.

TABLE III.-COMPARATIVE STABILITY DATA FOR THREE Pb_{1-x}Sn_xSe
LASERS STORED FOR OVER ONE YEAR AT ROOM TEMPERATURE.

Laser	Test Date	R _f (ohms)	I _{th} (mA)	ΔP	P _{max}	Laser History*
8324-10	11-20-78	.007	>2000	--	--	Control sample - No special cycling.
	12-07-78	.006	1380	43	78	
	12-18-78	.006	1350	71	109	
	1-08-79	.006	1280	58	93	
	2-08-79	.005	1330	59	102	
	2-24-79	.006	1350	69	104	
	5-25-79	.006	1500	54	90	
	11-06-79	.006	1330	49	97	
	2-20-80	.006	1400	40	90	
8324-11	11-20-78	.005	1490	50	84	Cycled 135 times Cycled 365 times 77-300K
	11-27-78	.005	1410	49	88	
	11-30-78	.006	1360	52	90	
	12-07-78	.005	1320	53	95	
	12-18-78	.006	1300	50	90	
	1-08-79	.006	1260	59	103	
	2-08-79	.005	1250	52	95	
	2-24-79	.006	1220	52	96	
	5-25-79	.005	1160	50	98	
	7-26-79	.006	1250	48	98	
	11-06-79	.006	1210	46	96	
	2-20-80	.006	1180	52	100	
8324-12	11-20-78	.006	1350	48	82	Cycled 135 times Cycled 365 times 77-300K
	11-28-78	.005	1270	43	85	
	11-30-78	.006	1200	51	96	
	12-07-78	.006	1200	52	95	
	12-18-78	.005	1250	51	91	
	1-08-79	.006	1220	50	96	
	2-08-79	.006	1230	48	90	
	2-24-79	.006	1260	42	86	
	5-25-79	.006	1160	43	94	
	7-27-79	.006	1360	35	76	
	11-06-79	.006	1400	23	65	
	2-20-80	.006	1310	30	70	

*All lasers stored in air at room temperature between tests.

TABLE IV.-STABILITY TEST DATA FOR LASERS 9093-21 AND 9110-3

Laser No.	Test Date	Electrical Contact Resistance (ohms)	Threshold Current (mA)
9093-21	4-04-79	0.007	+330
	4-09-79	0.008	+310
	4-23-79	0.008	+320
	7-13-79	0.007	+315
	1-11-80	0.004	+300
	2-12-80	0.005	+305
9110-3	4-20-79	0.007	-430
	7-30-79	0.007	-395
	8-31-79	0.006	-400
	2-12-80	0.005	-430

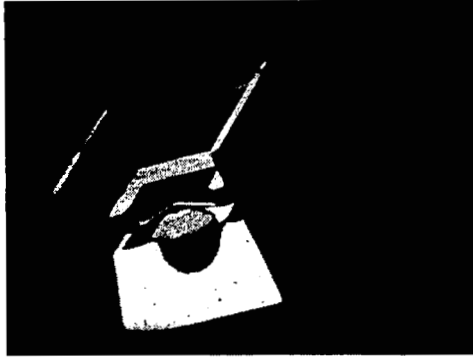


Figure 1.- SEM photograph of a Pb-salt diode laser package. This design minimizes crystal stress by eliminating pressure.

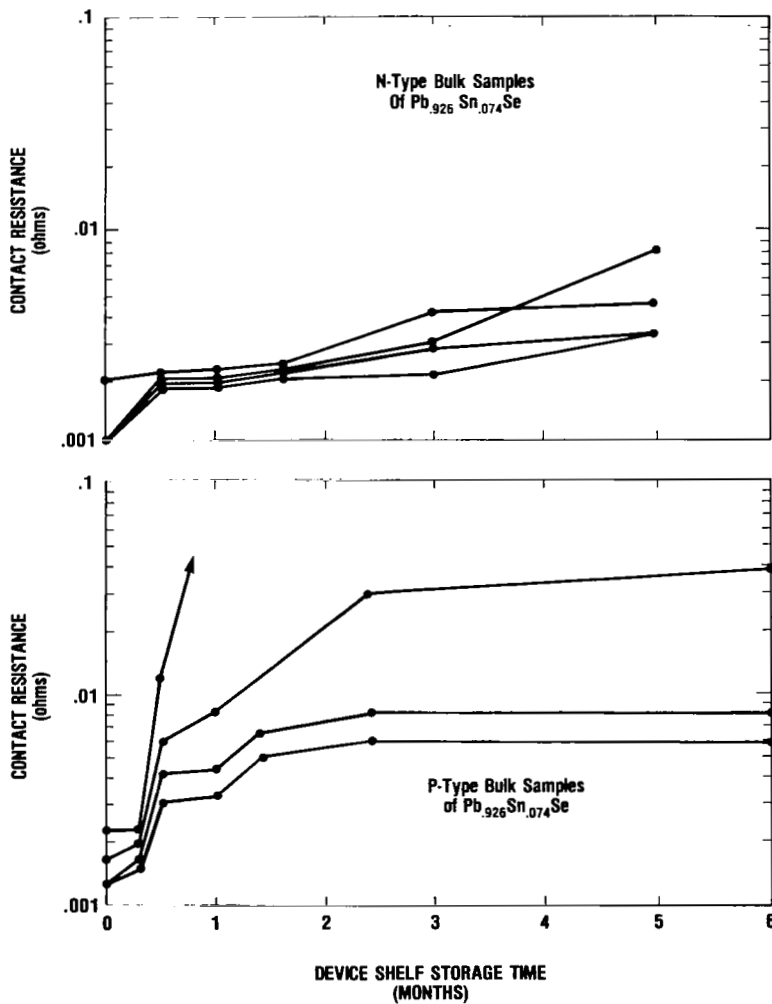


Figure 2.- Contact resistance to n-type and p-type $\text{Pb}_{0.924}\text{Sn}_{0.076}\text{Se}$ material with etched surfaces and standard contacts showing severe degradation.

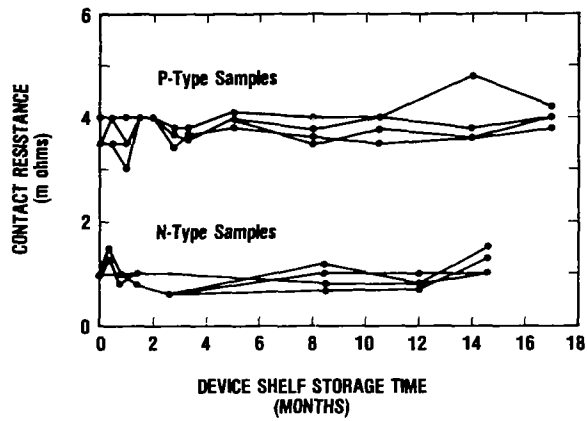


Figure 3.- Contact resistance to n-type and p-type $\text{Pb}_{.924}\text{Sn}_{.076}\text{Se}$ with improved contacts.

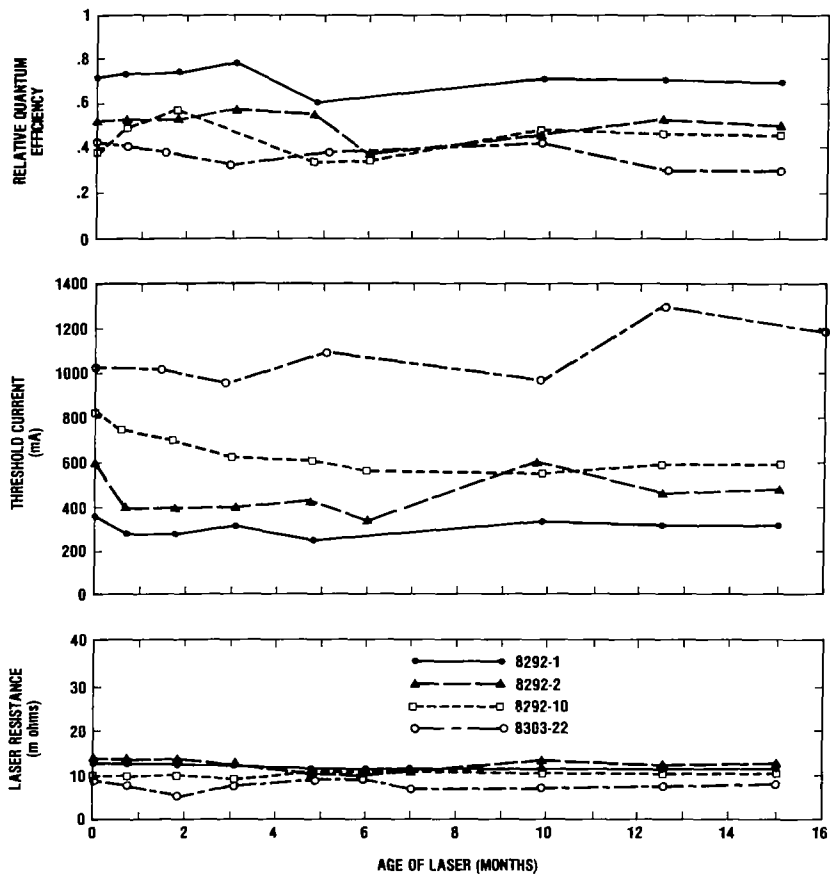


Figure 4.- Plots of relative quantum efficiency, threshold current and laser contact resistance as functions of shelf storage time covering a period of over one year. Lasers 8292-1 and 2 were fabricated from $\text{Pb}_{.966}\text{Sn}_{.034}\text{Se}$ and lasers 8292-10 and 8303-22 were fabricated from $\text{Pb}_{.927}\text{Sn}_{.073}\text{Se}$ corresponding to the 11.5 and 17.2 μm spectral regions, respectively.

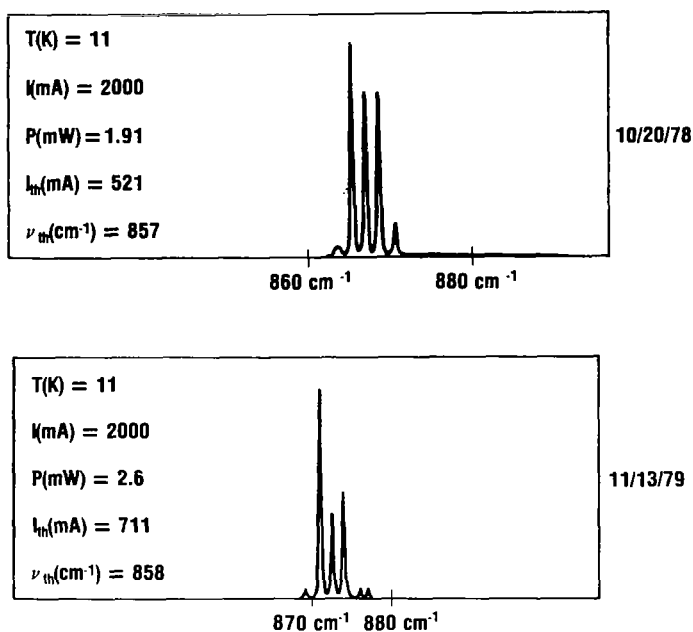


Figure 5.- Comparison of emission spectra from laser 8292-2 taken on October 20, 1978 and again on November 13, 1979. Differences in power reading could, in part, be due to different power meters. Shift of 6 cm^{-1} towards higher wave number is possibly due to slight thermal resistance increase.

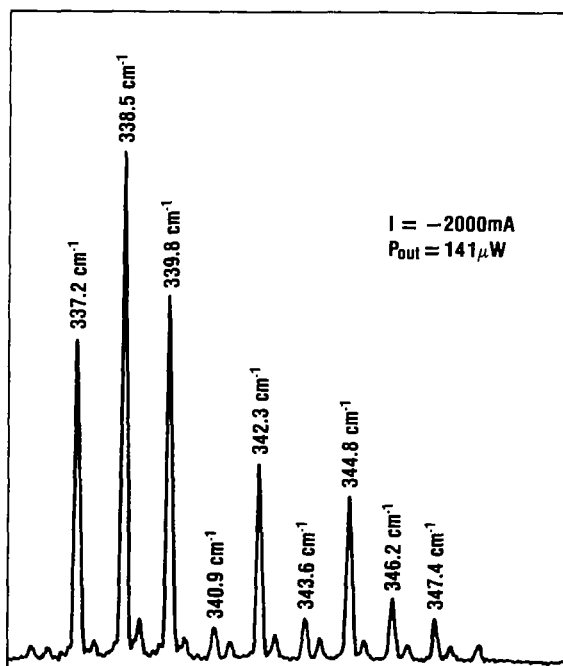


Figure 6.- Emission spectrum of laser 9096-10 operating CW at a 2 Amp current at 10 K. The 338.5 cm^{-1} mode has approximately $30\mu W$ of power.

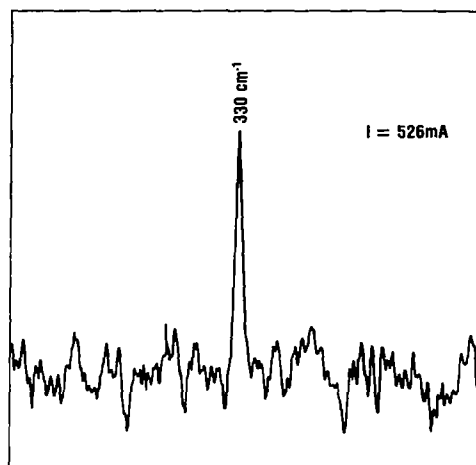


Figure 7.- Emission spectrum of laser 9096-10 operating slightly above threshold at 10 K showing single mode emission at 330 cm^{-1} ($30.30\text{ }\mu\text{m}$).

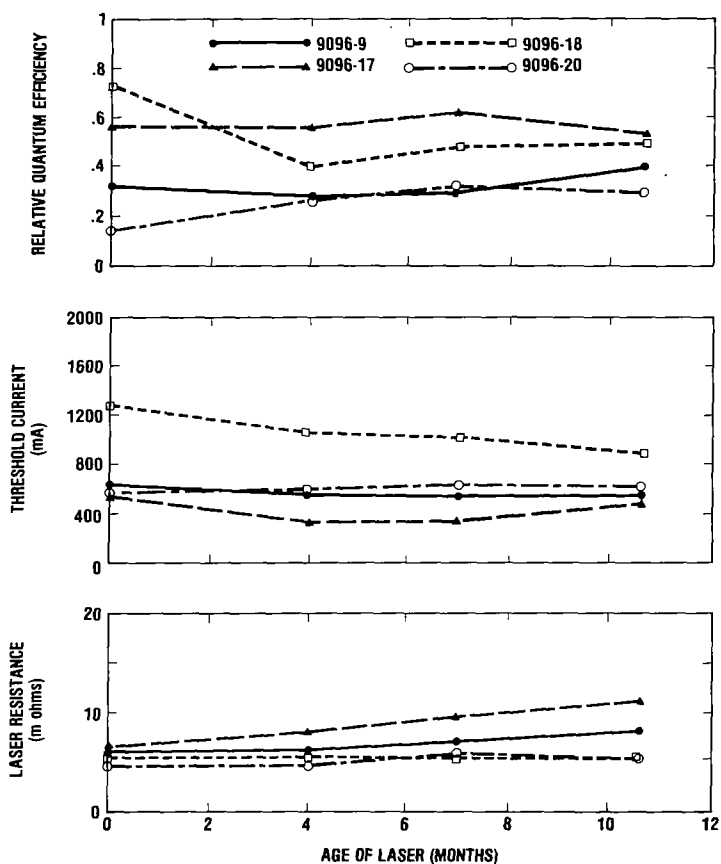


Figure 8.- Stability data for long wavelength ($28\text{ }\mu\text{m}$ region) $\text{Pb}_{0.90}\text{Sn}_{0.10}\text{Se}$ lasers stored under ambient conditions for a $10\frac{1}{2}$ month period.

LONG WAVELENGTH $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ AND $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ DIODE LASERS*

AS LOCAL OSCILLATORS

E. R. Washwell
Lockheed Palo Alto Research Laboratory

SUMMARY

A breadboard heterodyne receiver is described which has been used to establish the characteristics of lead salt diode lasers pertinent to their use as local oscillators using laser devices operating in the 15-25 μm spectral region. Heterodyne detection efficiency has been directly correlated with the transverse mode structure and emphasizes the necessity of stable lowest order mode operation for lasers when used as local oscillators. The results obtained indicate that the continued development of these lasers will provide suitable local oscillators for a variety of applications.

INTRODUCTION

The ability to operate $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ and $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ diode lasers at any given frequency over a wide wavelength range (approximately 7-30 μm) with a narrow linewidth has resulted in their application to high resolution infrared absorption spectroscopy. These same features provide the basis for their use as local oscillators in heterodyne receivers. However, additional characteristics are required of lasers used as local oscillators which are relatively unimportant in the usual incoherent (direct) detection mode of operation employed in absorption spectroscopy. In particular, the transverse mode structure and mode stability play a key role in the ability of the laser to provide quantum noise limited performance. Diode lasers which operate in the lowest order $\text{TE}_{0,0}$ mode provide excellent heterodyne mixing efficiency while those operating in higher order modes are typically inefficient and frequently unstable. These characteristics are of prime importance for heterodyne detection but have little influence on direct detection.

In the following sections, the laser power and mixer characteristics required for optimum heterodyne receiver sensitivity are summarized. The breadboard heterodyne receiver developed for the operation and characterization of lead-salt diode lasers as local oscillators is described in detail and results obtained on commercially available diode lasers presented.

* This work was supported by the Ballistic Missile Defense Advanced Technology Center under Contract DASG 60-76-C-0061

HETERODYNE SIGNAL-TO-NOISE RATIO

The current signal-to-noise ratio of an optical heterodyne receiver including polarization loss is given by (Ref. 1)

$$\frac{S}{N} = \frac{\left(\frac{\eta e}{h\nu}\right)^2 P_S P_{LO}}{\alpha e \frac{\eta e}{h\nu} P_{LO} B_1 (1 + \eta n_b) + \frac{4kT_N B_1}{G^2 R_A}} \left(\frac{B_1}{B_2}\right)^{\frac{1}{2}} \quad (1)$$

The numerator is the rms current resulting from the mixing of the signal power, P_S , with the local oscillator power, P_{LO} , followed by second detection. B_1 is the bandwidth following the first detector (IF) and B_2 that following the second detector; their ratio represents the signal-to-noise ratio improvement possible through the use of second detection with integration. The denominator contains the shot noise of the local oscillator and a thermal noise term, α equals 2 for a photodiode and 4 for a photoconductor, G is unity for a photodiode and is the photoconductive gain for a photoconductor, η is the mixer quantum efficiency, T_N and R_A the IF amplifier noise temperature and input impedance, respectively, e the electronic charge, k Boltzman's constant, h Planck's constant, and ν the frequency of the radiation. The term n_b is the number of background photons which are in a single spatial mode and within the IF bandwidth B_1 (and which therefore may interact coherently with the local oscillator), and is equal to $(\exp(h\nu/kT) - 1)^{-1}$ where T is the temperature of the background blackbody radiation. Equation (1) considers local oscillator shot noise (only) but does not include any excess noise, and assumes perfect wave-front matching of the signal and local oscillator beams at the mixer. The signal-to-noise degradation resulting from a wave-front mismatch has been treated elsewhere (Reference 2) and is not considered further here. It can be seen that in order to achieve quantum noise-limited operation with a heterodyne receiver, using an idealized local oscillator, the shot noise generated in the mixer by the local oscillator must exceed the thermal noise in the following amplifier. The local oscillator power P'_{LO} required to make these noise contributions equal is

$$P'_{LO} = \frac{4k}{e} \frac{h\nu}{\eta e} \frac{T_N}{R_A} \frac{(1 + \omega^2 \tau^2)}{\alpha G_O^2} \quad (2)$$

The power requirement for the laser is a function of the mixer element characteristics and the application, in particular, the IF bandwidth requirement. Diode lasers developed to date have been relatively low power devices ($\ll 1$ mW per mode), necessitating optimization in mixer and amplifier selection for a given application. For a variable local oscillator power P_{LO} , the receiver will approach quantum noise limited operation as $(1 + P'_{LO}/P_{LO})$. The actual signal-to-noise ratio achieved depends largely on such inherent laser characteristics as excess noise (mode instability) and transverse mode structure (mixing efficiency) and operational factors such as temperature stability and vibration

(shock) isolation, both of which affect the frequency stability. The low diode laser powers available in a single longitudinal mode during the course of this investigation dictated the use of photoconductive mixer elements and relatively low IF bandwidths. Photoconductive HgCdTe mixers were used between 10 μm and 17 μm and Ge:Cu for the longer wavelengths.

DIODE LASER HETERODYNE RECEIVER BREADBOARD

The laser characterization and heterodyne measurements were obtained using the apparatus shown in Figure 1. This system was used for the evaluation of diode lasers produced at Laser Analytics, Inc. (Bedford, MA) and New England Research Center (Sudbury, MA) operating in the 10-25 μm spectral region. The apparatus was also used for blackbody and molecular line radiation heterodyne measurements, and heterodyne detection of semiconductor laser emission at 23 μm and 10.6 μm .

The optics are all reflective (except for KRS-5 lenses in the frequency control loop and in the near-field measurement optical legs) which permits the receiver to be used over a wide wavelength range without refocusing, including the visible for alignment. The entire optical path can be enclosed with a plexiglass cover (not shown) and the system purged with dry nitrogen to eliminate atmospheric absorption. The lasers examined did not exhibit high beam divergence in a given single mode. This is as expected because of the long diffusion lengths and high refractive index in the Pb salts. Thus, the f/2 collection optics were more than adequate. The optical transfer system shown in Figure 1 matches the f/8 monochromator optics. The monochromator is used for spectral analysis and longitudinal mode isolation. The grating in the monochromator can be easily bypassed by a mirror which is used mainly for initial laser alignment. The radiation from the monochromator is collimated by a 4-in. focal length, off-axis, parabolic mirror to reduce the size requirement on the beam splitter. The beam splitter for the 15- μm spectral region consists of ZnSe 50% reflecting on one surface (at 45° angle of incidence) and antireflection coated on the other which is optimum for the weak signal case. Far-field radiation measurements are made in this collimated beam using a small area HgCdTe or pyroelectric detector mounted on an automated x-y stage which samples the cross section of the beam intensity profile. (Far-field measurements were also made at a distance of several centimeters from the diode laser, using no additional optics.) The X10 beam expander is employed to provide a convenient working distance to the focal point of the f/2 focusing mirror. The laser radiation transmitted by the beam splitter is passed through a 1-in. long solid germanium etalon (when used) for accurate tuning rate measurements and the radiation is also passed through a gas absorption cell which is used for accurate laser frequency measurements (referenced to a known absorption line) and for high-resolution spectroscopic measurements (including pressure broadening studies). This signal path also forms one leg of the frequency control loop which will be described below. The signal leg of the receiver can employ a blackbody source (operable in the Dicke switch mode), a heated gas emission cell (which can be operated in the heterodyne absorption mode by use of a blackbody source behind the cell), or a second laser at the positions shown in Figure 1. The emission cell was designed to have a low emissivity. Near-field (mirror radiation) and polarization measurements

have also been made as shown.

The major electronic components of the receiver are the laser driver, mixer, IF amplifier chain, and the frequency stabilization control loop. The components and circuits have been custom designed when necessary. All other components involve commercial instrumentation (IF amplifier, spectrum analyzer, lock-in detectors, etc.). The laser driver was designed for low-noise, low-drift output. A measured noise level of 20 μ A rms was achieved. Two drivers were built: a manually operated unit for general utility and a remotely controlled unit for use with the stabilization control loop. The Amplica 201 VSL IF amplifier used was thoroughly characterized having 5- to 110-MHz bandwidth, 1.2-dB noise figure, and 50-dB gain. This amplifier was used with the wider bandwidth Ge:Cu mixer. A Plessey SL1205 low noise preamplifier was used with the HgCdTe mixer. This amplifier was specifically designed for use with a HgCdTe detector having a 50-dB gain (nominal), a 6.5-MHz bandwidth and a $0.8 \text{ nV}/\sqrt{\text{Hz}}$ equivalent input noise voltage. In the direct detection mode, the output of the detector (mixer) circuit is connected directly to a lock-in amplifier and is used to monitor the laser mode power. In the heterodyne mode, the output of the IF amplifier is passed through a second detector which is referenced to the Dicke switch. The output of the IF amplifier can be directly connected to a spectrum analyzer for signal frequency analysis.

The mixers used in the 15 μ m spectral region were photoconductive HgCdTe 75 μ m and 150 μ m square elements mounted on a common heat sink. The peak spectral responsivity occurs at 14.3 μ m with good response to 16.5 μ m. The 40 nsec response time measured by the manufacturer (SBRC) using an InAs emitter was verified by the generation-recombination (g-r) noise roll-off frequency displayed on a spectrum analyzer and shown in Figure 2. This also demonstrates that the particular diode laser used was capable of generating a g-r noise level that exceeded the amplifier noise. The lifetime of this particular detector was the shortest available in high quality material. It may be possible to reduce the lifetime by compensation to provide a mixer with a wider bandwidth and flat response to higher frequencies, but this is advantageous only when increased laser powers become available. The breakdown field was measured to be 40V/cm which is sufficiently high to ensure attainment of the saturated drift velocity conditions at operational electric fields. The detector properties as a function of bias and background flux as well as responsivity were used to calculate the quantum efficiency ($\eta = 0.6$). The detector impedance is approximately 30 Ω (depending upon bias) and is relatively well matched to the IF amplifier.

The mixer element used for the 23- μ m spectral region was Ge:Cu. The lifetime and quantum efficiency were determined from measurements of the responsivity, D^* , and detector resistance at known background flux levels and as a function of bias field. This particular element had an interelectrode spacing of 0.0152 cm, a quantum efficiency of 0.15, a lifetime of approximately 5 nsec, and a saturated drift velocity of 5×10^6 cm/sec, resulting in a photoconductive gain of 1.64. The detector impedance at the flux levels during the experiments was approximately 1 M Ω . This high impedance level resulted in poor coupling to the 50- Ω IF amplifier used (Amplica 201 VSL) and nonlinear response over the amplifier bandwidth. The best 22 μ m laser did, however, induce

sufficient shot noise to exceed the amplifier noise, as observed on a spectrum analyzer, and provided a good blackbody heterodyne signal-to-noise ratio.

FREQUENCY STABILIZATION AND CONTROL

A significant advantage of a diode laser LO heterodyne receiver over fixed frequency laser systems is the capability to be operated at any frequency in the 3- to 30- μ m spectral region. Laser operation at a specific frequency is principally a function of the laser chemical composition, temperature, and drive current. A diode laser can be easily tuned (continuously over a 30-GHz range, for example) by simply changing the laser drive current. This versatility of diode lasers has a price, especially when used in a heterodyne application for the detection of molecular line radiation. Since the diode lasers are relatively low-power devices and efficient long wavelength infrared mixers presently have modest bandwidths, the laser must be operated close in frequency to the source of interest (within a few hundred MHz, for example) and must be stable to within a few MHz. This requirement makes frequency stabilization and control an important issue. Once the composition and nominal operating conditions have been established for a given device, the problem reduces to operation at a specific frequency with a given frequency stability. Since the laser tuning rates are typically 100 MHz/mA and 100 MHz/mdeg, close control of these parameters is essential. In principal, the absolute laser frequency can be determined by mixing the diode laser frequency with another fixed-frequency laser or harmonics of microwave devices in point contact devices. This technique is considered to be complex, inefficient, and unnecessary for many applications. The use of a secondary frequency standard is considered to be adequate. This standard can be simply a vapor phase molecular absorption which is at (or near) the line of interest. This absorption need not even occur in the same molecular specie as the line radiation of interest, but must be able to be established (isolated) to the required accuracy. Given such an absorption, the frequency stabilization technique investigated during this program involved operation of the desired laser frequency on the edge of the absorption line and use of an electronic feedback control of the laser drive current to maintain the laser output at the desired frequency.

A laser frequency stabilization and control loop was designed, fabricated, and tested. Two techniques were investigated; both are based on the stabilization of the laser frequency on the slope of a reference molecular absorption line. The first control loop utilized a single detector; the second used two detectors. The single mode selected by the monochromator is passed through an absorption cell whose line width can be pressure broadened from the Doppler limit to a few GHz. With the laser diode tuned to a frequency approximately half-way down the absorption line edge, the output voltage of a detector sensing this power transmitted through the absorption cell will increase or decrease proportionately with diode frequency, provided that there is no change in the laser power. The change in detector output is fed back through a controller to change the diode laser current and, therefore, its frequency. This will return the diode frequency to the original value selected on the absorption line. Fluctuations in diode power output will cause errors in the frequency change measurement and an independent measurement of power must be made.

The electronic system based on the single detector technique was bench tested to ensure proper operation of the controller. At this stage of the program, the experimental effort concentrated on the 23- μ m spectral region. The frequency control loop was, therefore, operated on the slope of an H₂O absorption line using a diode laser which could be tuned to the absorption line (although the laser power was relatively low at this particular frequency). The performance of the control loop in terms of frequency stability could not be satisfactorily evaluated due to large detector amplitude fluctuations caused by the reflective chopper wheel used in this mode of operation. The metallic wheel was found to have several distorted blades which caused the focused image at the detector to be displaced on and off the detector element. The resulting noise level of the system was too large for effective operation of the control loop. A second system was constructed which made use of two detectors rather than one. The basic principle is the same as for the case of a single detector; the difference is in the complexity of the detailed electronics. The loop was tested electronically but not operated with a laser.

It should be pointed out that considerable effort has been applied to the passive frequency stabilization of the laser output. For example, the laser current source (drive) provides a low-noise, low-drift output and the liquid helium Dewars used for laser cooling have been superinsulated to minimize varying external heat loads. The mechanical refrigerator (Air Products Model CS 202-modified) has a specially designed thermal dampener to reduce temperature fluctuations caused by the piston cycle. Indirect frequency stability measurements of the laser output in the closed cycle refrigerator show that frequency variations greater than a few MHz did not occur in general, although significant laser amplitude (power) fluctuations at the mixer can occur for a poorly mounted refrigerator in addition to excess laser noise induced by the mechanical refrigerator. (Fig. 3 shows a Doppler limited line width was achieved in CO₂.) These amplitude fluctuations were the result of a physical displacement of the laser and its conjugate image at the small mixer element used and were minimized by proper mounting arrangements. In all the measurements to be reported, the only active laser frequency control loop involved temperature stabilization of the mechanical refrigerator.

A more direct measure of the frequency stability attainable using these techniques is shown in Fig. 4. The heterodyne beat note obtained using a CO₂ laser as the local oscillator and a diode laser operating in the modified refrigerator as the signal source is shown. The line width remained unchanged when the refrigerator was turned off and the line position swept thru the IF bandpass as expected. The observed line width of approximately 5 MHz indicates excellent frequency stability.

MECHANICAL REFRIGERATOR MODIFICATION

The manufacture of a diode laser operating at a specific frequency is a time-consuming process. Crystal growth, annealing, junction formation, contacting, and yield limit the number of devices that can meet a given specification. In order to increase the usefulness of a given device's chemical composition and, therefore, reduce the fabrication complexity, lasers are

operated in mechanical refrigerators which provide additional convenient (temperature) tuning capability. The use of such a cooler, however, can introduce additional unwanted mechanically induced laser noise. Quantitative measurements were made of the shock and vibration effects in the Air Products refrigerator. The displacement of the cold head due to the piston motion was measured to be 0.015 cm (0.006 in) using a capacitance technique. Shock effects were investigated using an accelerometer attached to the cold head. The data were acquired and analyzed by a Hewlett-Packard 5450A Fourier Analyzer. A typical spectrum and its Fourier transform are shown in Figures 5 and 6 respectively. The analysis shows a relatively complex spectrum having a peak value of approximately 8g with a major component at 1500 Hz. The results of these data were used in the shock/vibration isolator design. This design is similar to that used by Dr. Jennings at NASA Goddard Space Flight Center (Ref. 3), but is more versatile in providing five degrees of freedom in the laser positioning, which we have found to be necessary to provide the proper alignment of the laser in the optical system.

The effectiveness of the shock isolation can be seen from a comparison of Figures 7 and 8. Identical accelerometers were placed on the cold head and the modified laser mount. Figure 7 shows the accelerometer output at the refrigerator cold head and Figure 8 the corresponding data at the modified laser mount. A dramatic decrease in the peak g-value is to be noted as well as a reduction in the high-frequency content. A total of six laser mounts were made to accommodate all presently existing laser package designs and comparative accelerometer data were obtained.

The laser measurements taken with the modified refrigerator included a comparison of the direct detected laser power level (in a given longitudinal mode), the second detection (IF noise) level, and low-temperature blackbody heterodyne SNR measurements when possible. The magnitude and noise level of the IF signal (and, therefore, heterodyne signal-to-noise ratio (SNR)) is a strong function of the longitudinal mode stability which depends upon the laser operating conditions. The SNR attainable for the heterodyne detection of thermal radiation can be directly related to the mode stability. Certain modes of a given laser are quite stable over a wide range of conditions and others are unstable under any condition. A comparison of the IF noise level produced by a laser when operated in the modified refrigerator and a liquid helium Dewar showed that the modification of the refrigerator greatly reduced the random (undesired) noise level but did not eliminate it; i.e., liquid cryogen cooling provides a better heterodyne SNR than mechanical refrigerator cooling. In any case, the modified refrigerator decreased the excess IF noise level to the extent that the inherent mode instability and poor transverse mode structure of the diode lasers become the major limitations in achieving near quantum-noise limited heterodyne performance.

LASER MEASUREMENTS

The diode lasers characterized and used during this program include lasers operating in the 23- and 15- μm spectral regions. Detailed measurements were made on more than 20 lasers. This section will report results obtained on selected devices which typify diode laser characteristics. The I-V characteristics of the lasers had series resistance (R_s) values from tens of milliohms to as high as 180 m Ω . Good I-V characteristics in general did not correlate with laser performance. As an example, lasers having the same R_s value would vary in output power by a factor of 10 and lasers exhibiting the well-known kink, supposedly indicating high radiative efficiency, frequently would have lower power than a laser with a more gradually rising I-V curve. Although a few lasers exhibited an extremely large number of longitudinal modes, a more typical behavior was three to six longitudinal modes operating simultaneously at current levels sufficiently above threshold to provide useful laser power levels. The typical continuous tuning range of a given mode for all lasers examined was approximately 1 cm^{-1} (30 GHz). The power output characteristic for a continuously tunable mode varied considerably. The output power peaked at any point in the range from lasing onset to the disappearance of the mode. The laser which gave the best blackbody heterodyne detection results showed efficient IF noise generation for all modes as shown in Fig. 9.

The beam divergence we have observed has varied from $f/1$ to $f/10$ with values larger than $f/2$ being more common. Some of the observed far-field patterns can be explained by a blockage (reflection) of a portion of the diverging beam by the heat sink, which is a fabrication problem and is not fundamental to the device. All the lasers examined to date are polarized in the plane of the junction and frequently in the lowest order $\text{TE}_{0,0}$ mode. The near-field (mirror radiation) pattern of two lasers selected at random showed only one active region and no evidence of filamentary action (which is understandable in light of the long diffusion lengths in the Pb salts).

The major emphasis to date in the development of the lead salt lasers has been concerned with increasing the total output power. Considerable success has been achieved in this area, particularly in the past two years. Output powers in the milliwatt range are common, compared with microwatt levels in the early stages of development. Although a sufficiently high power level is a necessary condition for a local oscillator source, it is not sufficient by itself. The power must be also available in a single longitudinal mode. Unfortunately, the presently available higher power lead-salt lasers exhibit a relatively large number of longitudinal modes, limiting the power available in a single mode. Furthermore, the laser emission radiation pattern frequently consists of several diverging lobes. The angular extent of a given lobe is relatively small as mentioned above, but the total angular extent may be large. This situation is shown in Fig. 10, which is the far-field pattern of a diode laser operating at 15.4 μm (cw) obtained from a cross sectional scan of the output beam without using additional optics. Three lobes are evident and each, in turn, could be isolated by the monochromator by using the angular adjustments available in the modified refrigerator laser mount. Although $f/1$ optics can collect all of this emitted radiation, the available power in a single lobe (mode) is limited.

The lowest order TE mode is required for good heterodyne mixing efficiency and therefore the importance of the transverse mode structure of the laser in determining the mixing efficiency cannot be overstated. A typical result is shown in Figure 11. It can be seen that individual longitudinal modes have varying degrees of mixing efficiency; poor efficiency is correlated with a complex transverse mode structure of an individual longitudinal mode. This particular laser also exhibited significant excess noise resulting in a poor heterodyne signal-to-noise ratio as compared to theoretical predictions. The far-field intensity distribution for a relatively efficient mode is shown in Figure 12(a) and in Figure 12(b) for an inefficient mode. The plots represent a partial cross-sectional scan of the laser beam using a small detector in the collimated beam of the laser. Close inspection of Figure 12(b) reveals a more complex intensity distribution than that in Figure 12(a). These results are general: higher order transverse modes result in poor heterodyne detection efficiency.

The transverse mode structure of a diode laser is only of minor importance in high-resolution spectroscopy and has consequently received little attention in the manufacture of these lasers. In general, to ensure lowest order mode operation, additional optical and carrier confinement is expected to be required over that occurring in the simple diffused junction devices presently manufactured, although homojunction devices which operate in the lowest order mode have been produced which provide excellent heterodyne mixing efficiency. The nature of the mode instability (excess noise) is not well understood at present. It is possible that improvements in material uniformity and crystal perfection will relieve this problem. Self-beating effects have also been observed with individual longitudinal modes and may be due to Fabry-Perot cavity imperfections produced during the cleaving procedure to form the cavity. These effects result in excess noise in the detection process but do not occur for the majority of the modes and can frequently be minimized by changes in the operating conditions.

These results emphasize the importance of the cavity configuration of diode lasers to be used as local oscillators. It is essential that the cavity provide a minimum in longitudinal mode competition as well as a strong rejection of higher order transverse modes.

SUMMARY

A breadboard heterodyne receiver is described which has been used to establish the characteristics of lead salt diode lasers pertinent to their use as local oscillators using laser devices operating in the 15-25 μm spectral region. Heterodyne detection efficiency has been directly correlated with the transverse mode structure and emphasizes the necessity of stable lowest order mode operation for lasers when used as local oscillators. The results obtained indicate that the continued development of these lasers will provide suitable local oscillators for a variety of applications.

REFERENCES

1. Washwell, E. R.; and Ichiki, S. K.: SPIE Proc. Vol. 95, pp 41-46 (1976).
2. Cohen, S. C.: Heterodyne Detection: phase front alignment, beam spot size, and detector uniformity. Appl. Opt. 14, 1953-1959 (1975)
3. Jennings, D. E.; and Hillman, J. J.: A Shock Isolator for Diode Laser Operation on a Closed-Cycle Refrigerator. Goddard Space Flight Center Document X-693-77-143, June 1977.

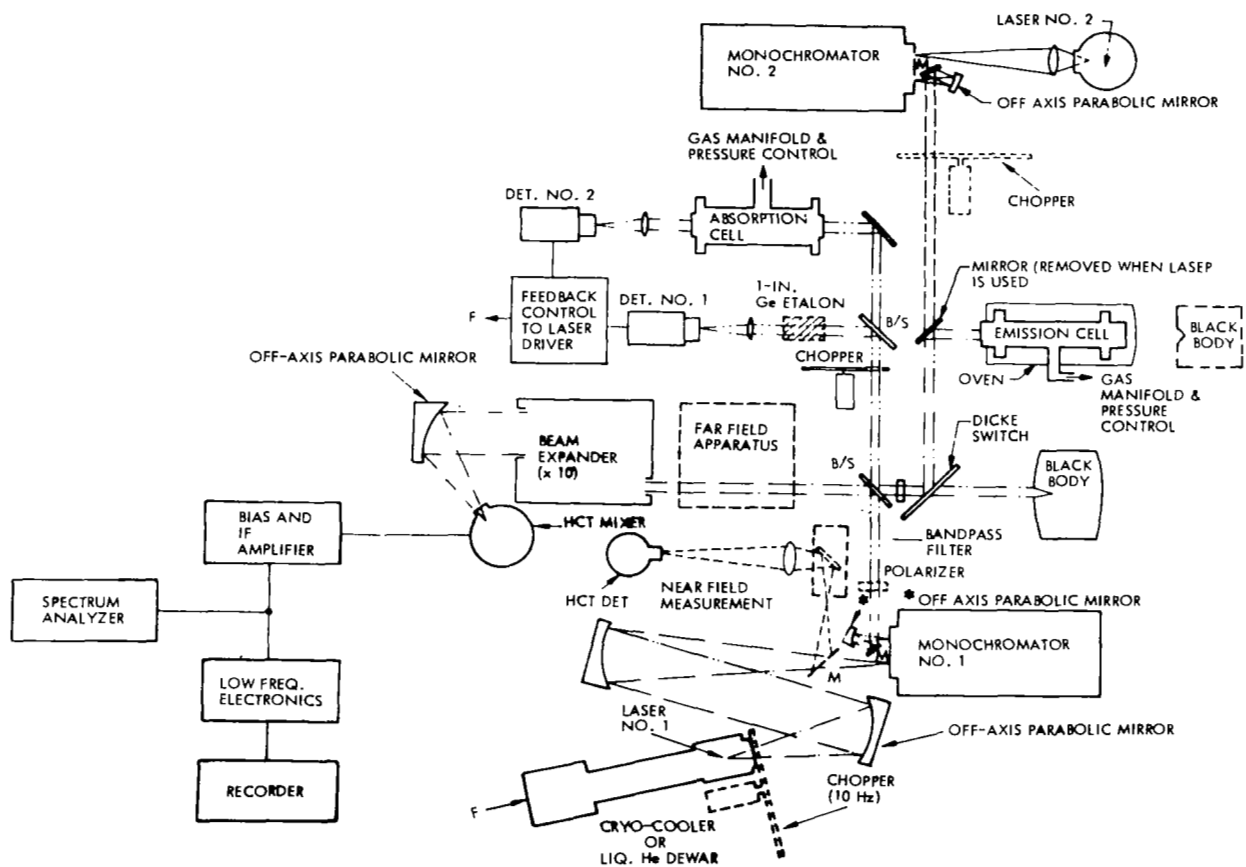
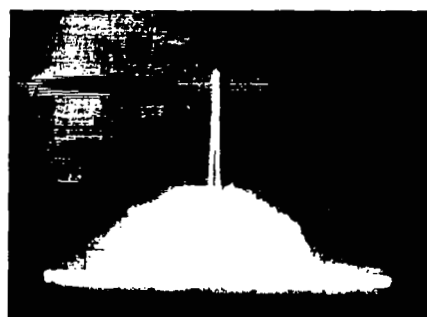


Figure 1.- Diode laser heterodyne receiver breadboard.



(A) LASER OFF
5-MHz BANDPASS
FILTER,
2 MHz/cm
HORIZONTAL
AXIS



(B) LASER ON
5-MHz BANDPASS
FILTER
2-MHz/cm
HORIZONTAL
AXIS

Figure 2.- Generation-recombination noise generated by 15- μ m diode laser in HgCdTe.

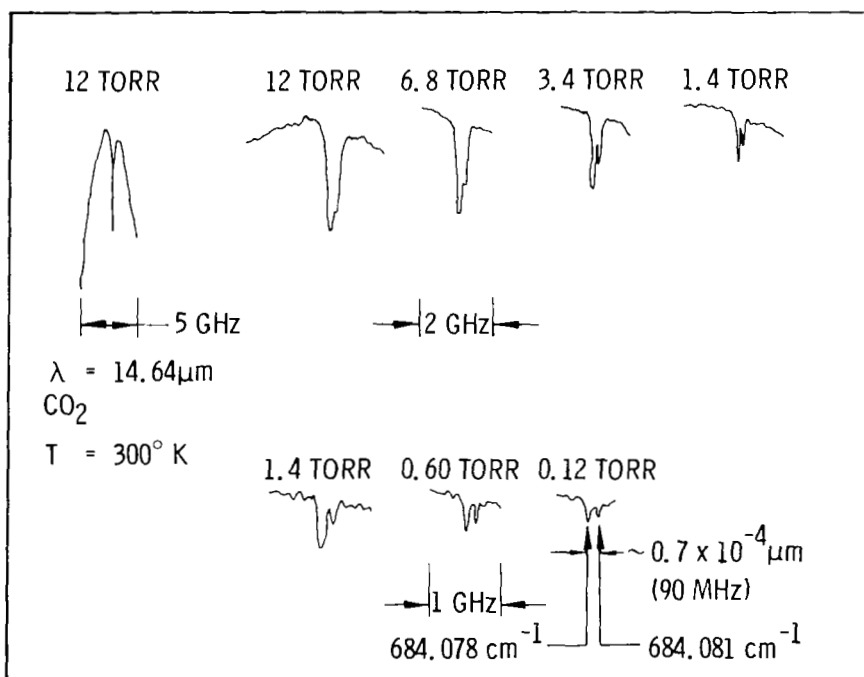
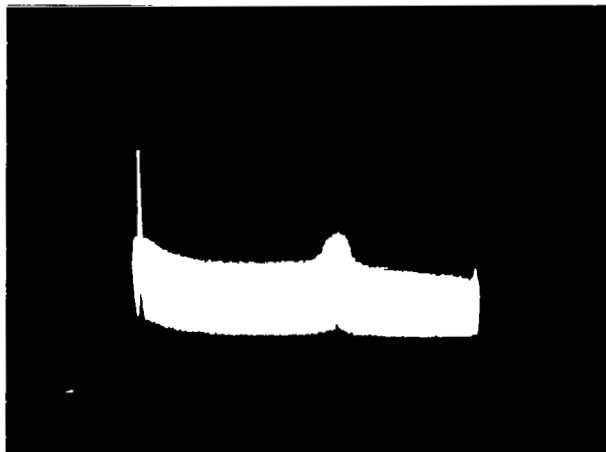


Figure 3.- Frequency stability measurement inferred from Doppler line width.



10 MHz PER DIVISION;
2-ms SWEEP

Figure 4.- Direct frequency stability measurement from heterodyne beat note.

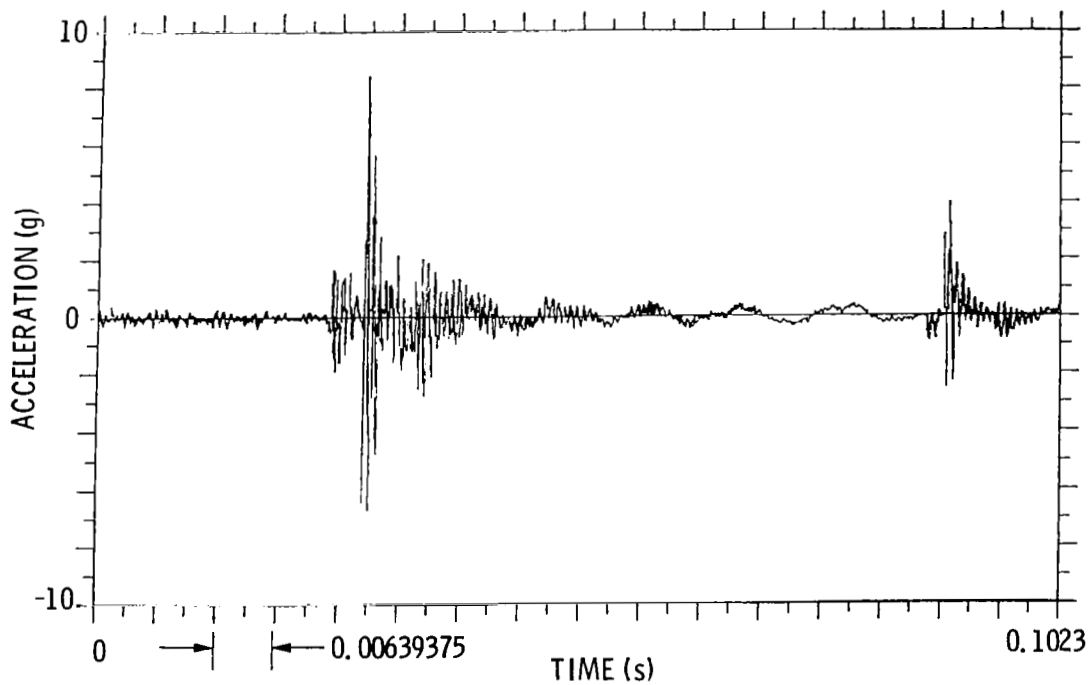


Figure 5.- Refrigerator accelerometer signal.

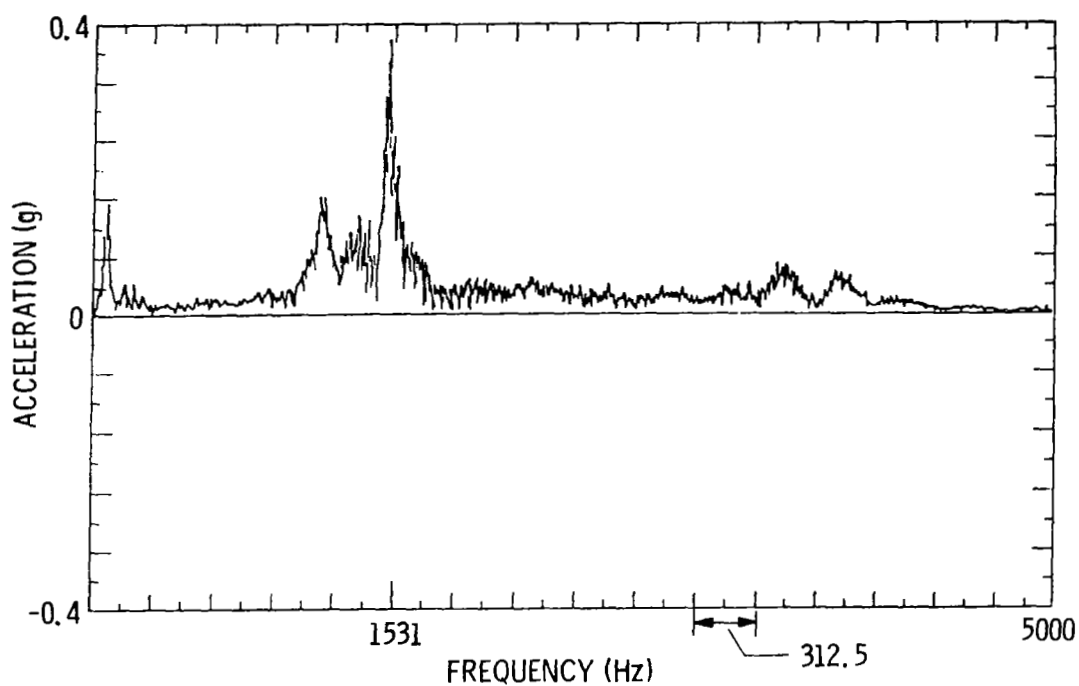


Figure 6.- Fourier transform of accelerometer output.

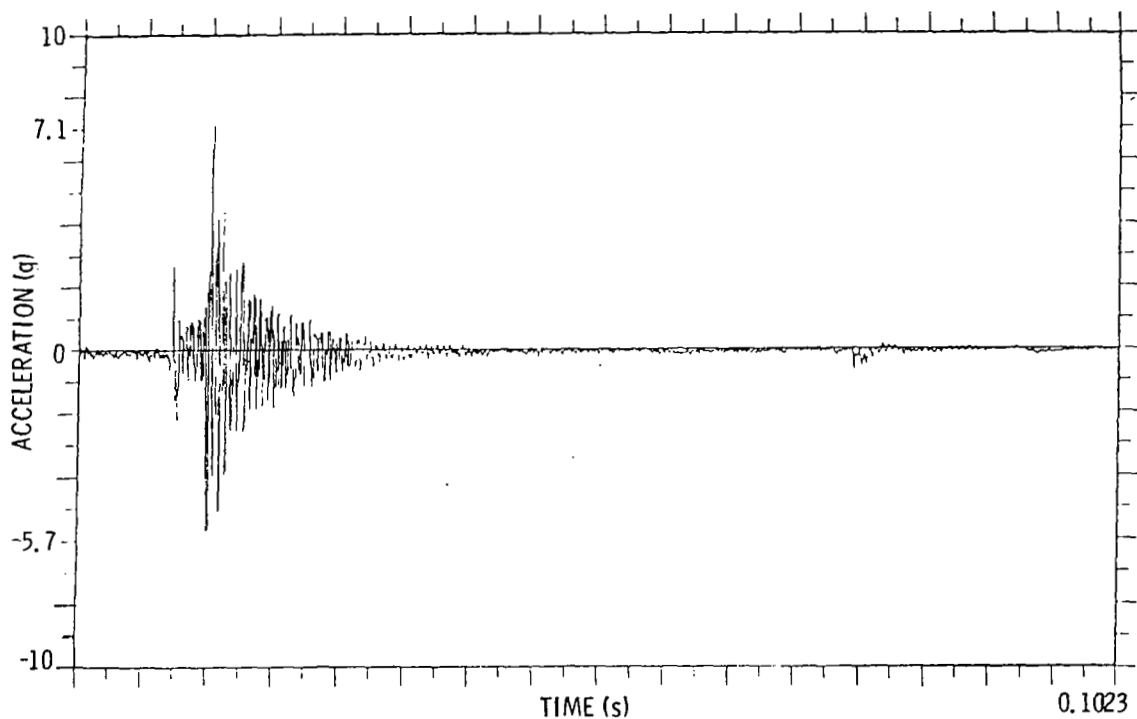


Figure 7.- Modified refrigerator accelerometer signal (at cold head).

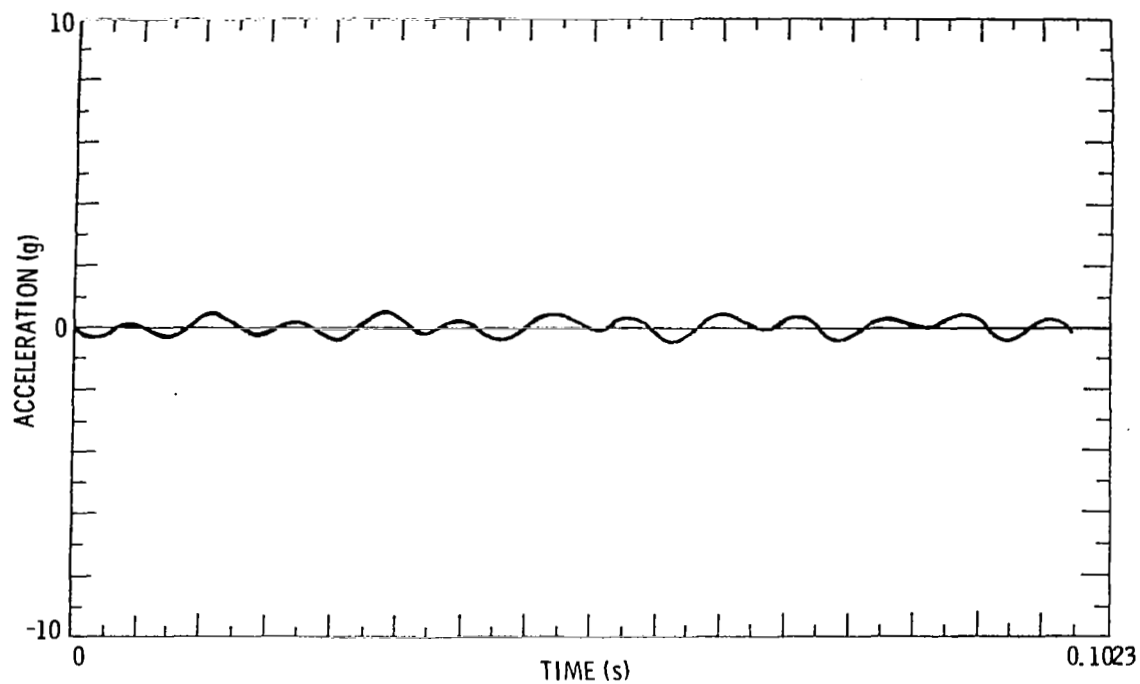


Figure 8.- Modified refrigerator accelerometer signal (at laser mount).

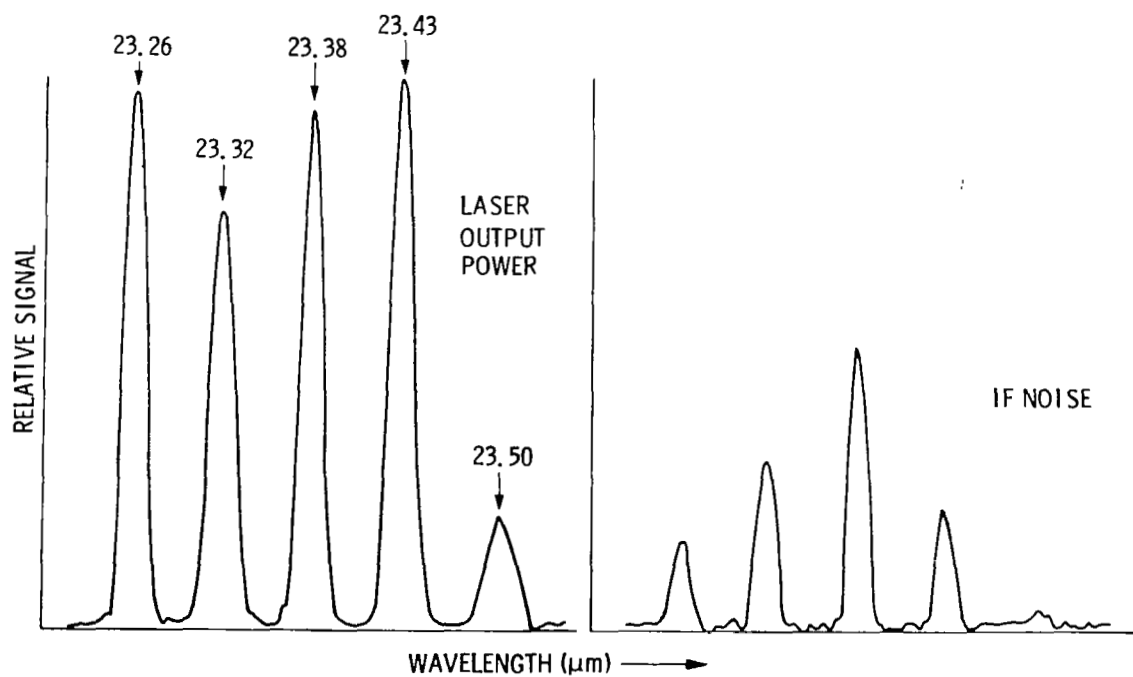


Figure 9.- Correlation of laser mode power and IF noise level.

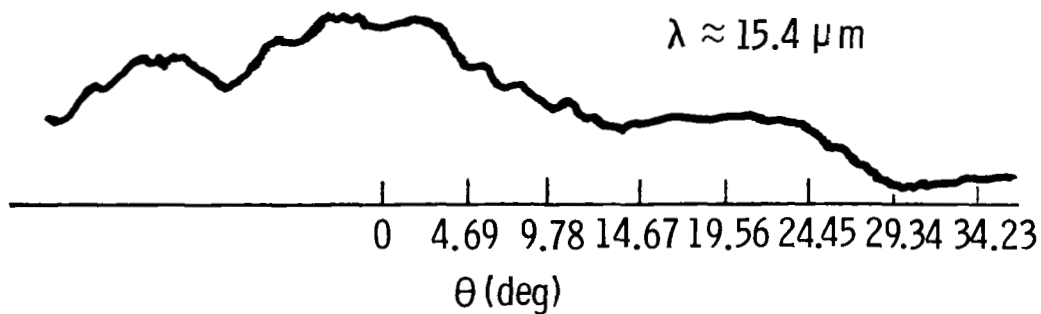


Figure 10.- Far-field beam profile (15- μm laser cw).

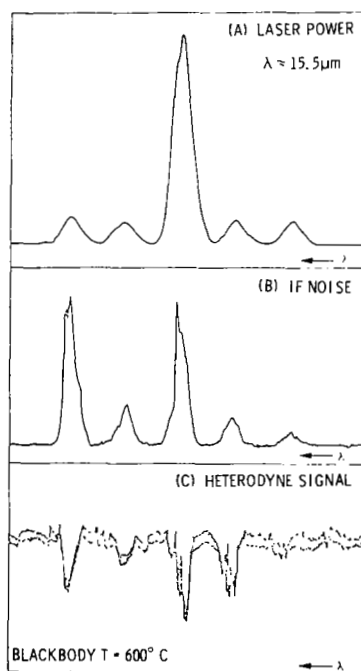
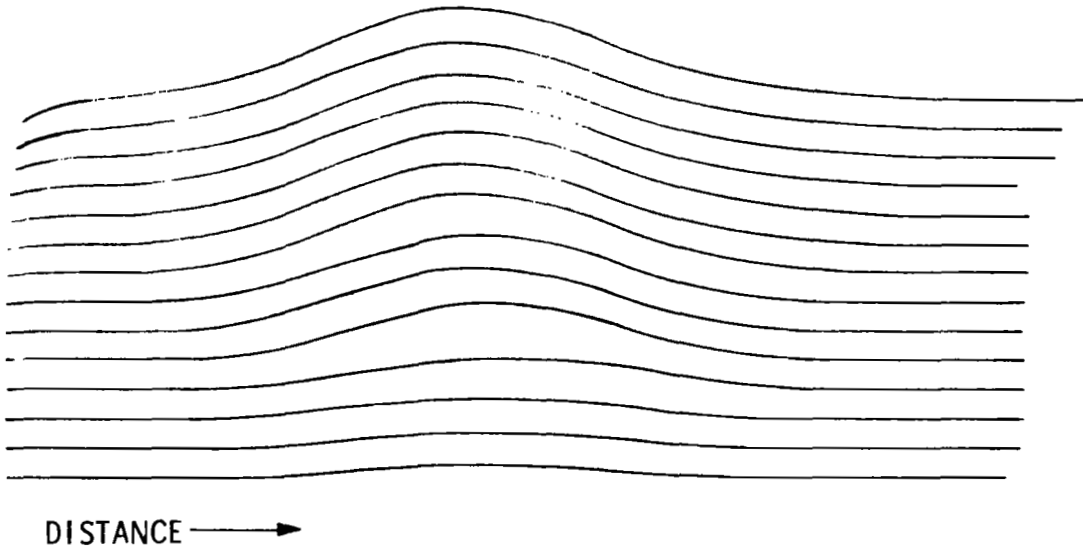
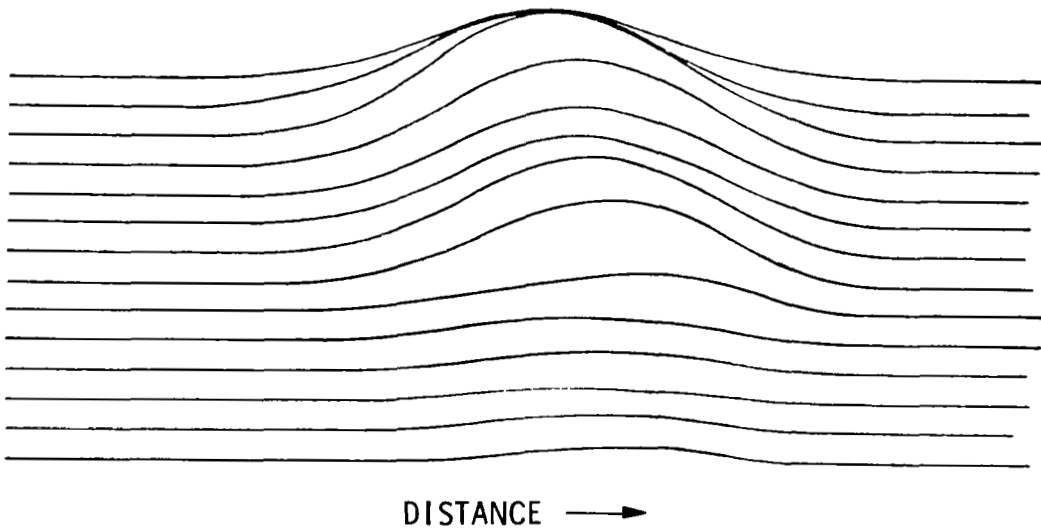


Figure 11.- Heterodyne spectrum as a function of laser power and IF noise level.

—| |— RESOLUTION



(a) Efficient mode.



(b) Inefficient mode.

Figure 12.- Laser far-field profile.



DIAGNOSTICS AND CONTROL OF WAVENUMBER STABILITY
AND PURITY OF TUNABLE DIODE LASERS RELEVANT TO
THEIR USE AS LOCAL OSCILLATORS IN HETERODYNE SYSTEMS*

S. Poultney, D. Chen, G. Steinberg, F. Wu, A. Pires, M. Miller,
and M. McNally
The Perkin-Elmer Corporation, Norwalk, CT. 06856

SUMMARY

Performance characteristics inherent to both available tunable diode lasers and their environment will affect their use as local oscillators in heterodyne systems. This paper reports diagnostics of wavenumber stability and purity of tunable diode lasers in closed cycle cooler environments, and a new concept for control of the wavenumber in heterodyne system applications. The inherent TDL characteristics included the capability to be tuned to the wavenumber desired, the change in operating point as a function of time and thermal cycling, the nonlinear dependence of wavenumber on drive current, the axial mode structure, and the observed fine structure of "single" axial modes.

Initial operation of the TDL showed that it was not possible to adjust the wavenumber to one selected a priori in the TDL tuning range. Its operating point was recorded as a function of time and thermal cycling by heterodyne and absorption techniques. During operation, the operating point would change by 0.1 cm^{-1} over the longer term with even larger changes occurring during some thermal cycles. Most changes during thermal cycling required using lower temperatures and higher currents to reach the former wavenumber (when it could be reached). In many cases, an operating point could be selected by changing TDL current and temperature to give both the desired wavenumber and most of the power in a single mode. The selection procedure had to be used after each thermal cycling. Wavenumber nonlinearities of about 10% over a 0.5-cm tuning range were observed. Diagnostics of the single mode selected by a grating monochromator showed wavenumber fine structure under certain operating conditions. The characteristics due to the TDL environment included short-term wavenumber stability, the instrument lineshape function, and intermediate-term wavenumber

* This work was partially supported by the High Altitude Pollution Program, Contract No. FAA-AEE-79. However, the primary sources of support were Perkin-Elmer's Independent Research and Development Programs and capital expenditures. S. Poultney wishes to thank Dr. David A. Huchital, Director of Research, for hosting the measurement programs and for having the foresight to provide the initial capital support.

stability. Steinberg (ref. 1) has reported wavenumber set-point stabilities of several MHz on the short term (up to several minutes). His heterodyne measurement of this stability indicated that the instrumental line-shape function had a half-width of 10 MHz and a shape with broad wings and that the wings were due to residual cooler vibrations which could not be removed. Preliminary work with digital control of TDL current indicates an intermediate-term wavenumber stability (up to an hour) of 10^{-3} cm^{-1} or better which is related to cycling of the temperature control system.

These wavenumber stability and purity characteristics will require supplementary systems to ensure that the TDL will be an appropriate local oscillator, especially if thermal cycling may occur. We have developed the following concept for wavenumber control such that the TDL wavenumber can be adjusted to and held at the desired wavenumber location in the absence of a calibration cell containing the gas to be measured. The wavenumber control algorithm is designed to initialize the TDL wavenumber to that of an absorption line of a control gas after thermal cycling, step the TDL wavenumber in the presence of scan nonlinearities to the desired wavenumber which is assumed to be within a mode tuning range, lock to that desired wavenumber in spite of intermediate and long-term wavenumber drifts, and monitor continually the stability of the stable etalon vernier. The stepping and locking is accomplished using fm/harmonic detection of etalon fringes in quadrature in an auxiliary absorption spectrometer branch. This technique allows use of moderate finesse fringe patterns through its unique filtering capabilities and is insensitive to variations in TDL power. Details of the software and hardware design will be given including the design of the stable etalon.

We conclude that the characteristics of available laser diodes and aspects of their performance in closed-cycle cooler environments which may have made space flight systems impractical can be overcome through realizable control techniques.

INTRODUCTION

In 1976, S.K. Poultney initiated a two-pronged research program to enable Perkin-Elmer to return to the flight spectroscopic sensor business. This program emphasized the FTS interferometer as a quantitative-survey instrument and the tunable laser spectrometer as a species-specific monitor. The tunable laser spectrometer was based on the Pb-salt tunable diode lasers. It encompassed the applications accessible to the Laser Heterodyne Spectrometer (LHS) and to the Tunable Diode Laser Absorption Spectrometer System (TDLSS). Following the assembly and testing of a CO₂-laser based LHS, S.K. Poultney led the assembly and development of the state-of-the-art TDLSS for trace gas detection.

The following section describes the Perkin-Elmer TDLSS. The trace gas detection measurements with this system have been supported by the Federal Aviation Administration; first as a feasibility study (ref. 2) and more recently as a flight prototype activity led by M.D. Miller. These measurements with the TDLSS as well as many performance diagnostics supported by independent research programs have allowed us to obtain the measurements that we report in this paper. The goals of our investigations of the tunable laser spectrometer have been to determine those characteristics of the spectrometer which will limit its use for either of its two applications and to develop a potential solution for one of the most critical performance aspects. That aspect of performance is wavenumber purity and stability. Table I summarizes the wavenumber purity and stability required by the most demanding application; a laser heterodyne spectrometer measurement from space of reactive trace gases which cannot be contained in a reference cell and whose absorption lines have Doppler widths. The next two sections report the results of our wavenumber purity and stability diagnostics for performance characteristics inherent to tunable diode lasers and their environment, respectively. The final section describes a new concept for automatic wavenumber control which has been developed on an independent research program, and briefly summarizes our design for the necessary stable Fabry-Perot interferometer. Methods for ensuring a single mode at the detector are not discussed.

DESCRIPTION OF THE TUNABLE DIODE LASER SPECTROMETRIC SYSTEM (TDLSS)

The tunable diode laser spectrometric system consists of a laser source, its control environment, a sample cell, a detector, connecting optics, and a data acquisition system as shown in figure 1. The laser must normally be kept at temperatures near 25 K with a precision of tenths of millidegrees. It is normally tuned in wavenumber by applying an incremental drive current to the necessary bias current from a control module. This particular system possesses a digital control box which steps the drive current in small steps and synchronously clocks the A/D converter in the data acquisition system. This TDLSS can be operated in either the conventional am detection mode using the chopper or the fm/harmonic mode using one or more fm modulations superposed on the drive current. The signal out of the detector is appropriately demodulated and filtered before being digitized. A number of auxiliary optical devices are needed to control the laser output and to calibrate the laser wavenumber. These devices include a mode selector, a coarse wavenumber locator, calibration gas cells, calibration etalons, feedback isolators, etc. For use as a laboratory spectroscopy system, the TDLSS requires all these components as well as an executive software program.

The tunable laser source is a semiconductor diode of the Pb-salt family which can emit substantial amounts of power in a narrow spectral range in the midinfrared under the proper conditions. All of the lasers that we have used have been purchased from one commercial source, Laser Analytics, Inc. These diodes have yielded about 100 microwatts in each of several modes in the wavenumber region specified. Their lifetimes have been excellent (e.g. many months). Our specified wavenumbers have been reached within the available ranges of closed-cycle cooler temperatures and control-module bias and drive currents.

The TDL's must be operated under the proper temperature, mechanical-mount, and current conditions to insure best performance. For the laboratory TDLSS we selected an AIRCO closed-cycle cooler to provide the necessary cryogenic environment. In this environment, the TDL can operate between about 20 K and 50 K depending on its composition and the wavenumber desired. Temperature control and vibration mounting of the TDL was engineered by Perkin-Elmer and is described by Steinberg (ref. 1).

The TDL requires a current bias of up to 2 A in order to emit light. It is normally tuned in wavenumber by applying a variable drive current. Typical current tuning coefficients are $0.011 \text{ cm}^{-1}/\text{mA}$. We selected the analog current-control module made by Laser Analytics Inc. for the Laboratory TDLSS. We have found this to be a well-engineered unit. We have also found it convenient to supplement the analog current-control module with a digital control box. This digital control box steps the drive current in small steps and synchronously clocks both a chart recorder and the A/D converter in the data acquisition system. The data acquisition system records the digitized TDLSS scans for later analysis on the Perkin-Elmer time-sharing computer. The digital control box can scan at different rates and different tuning ranges, can scan down, can digitally set the bias current, and can accept either wavenumber control signals or several TDL modulation signals at an external summing junction.

This TDLSS can be operated in either of two detection modes: am chopping or fm/harmonic detection. The conventional am mode uses the chopper to modulate the light beam and the synchronous amplifier to demodulate and filter the detector signal. We use this mode for spectroscopic measurements of line positions, strengths, and widths (ref. 3). The fm/harmonic mode uses one or more fm modulations superposed on the TDL drive current and the synchronous amplifier to demodulate and filter the detector signal. We use this mode for trace gas detection because of its advantages for this application as discussed by Poultney et al. (ref. 2).

The detector is a broad response HgCdTe detector from Hughes/SBRC. It was purchased in conjunction with a matched, low noise preamplifier. The detector has performed well, but it was not specifically matched for any single application. Neither the field of view, the size, nor a cold filter has been selected to provide the lowest noise achievable in the present TDLSS.

The TDLSS optics consist entirely of offaxis parabolas and plano mirrors in order to minimize parasitic etalon noise and to make alignment straightforward. The optics are modular in form to allow convenient change of functions.

A number of auxiliary optical devices are needed to control the laser output and to calibrate the laser wavenumber. A grating monochromator (1/2 m Jarrell-Asch) has been used to locate the laser wavenumber (to about 0.5 cm^{-1}), to study the axial mode structure of the TDL, and to select a single mode for transmission through the TDLSS. The monochromator has worked well in a laboratory environment and is essential for the TDL diagnostics and coarse wavenumber adjustment. The absolute wavenumber is best determined by the use of a gas with known absorption lines in a calibration cell.

For use as a laboratory spectroscopy system, the TDLSS requires an executive spectroscopy program which can read the digital records, normalize and calibrate each record, connect each short record (e.g. 1 cm^{-1}) into the more normal spectroscopic display, print out the spectroscopic display, utilize analysis subprograms for the analysis of the data, and archive the data and results of the analysis. The flow chart of that executive spectroscopy program is shown in figure 2. Some of the analysis programs are library programs while others have been specifically written to accomplish nonlinear, least square fitting of spectral lineshapes (ref. 4).

DIAGNOSTICS OF INHERENT TDL CHARACTERISTICS

Capability to be Tuned to Desired Wavenumber

Our initial work with the NO TDL showed that we could not obtain an operating point which gave an output at the R(4.5) NO lines, but could obtain a good output at the nearby R(5.5) NO lines. This wavenumber search is executed by making a series of changes in both TDL heat sink temperature and drive current. The drive current tunes the TDL through joule heating in a finer degree than the temperature controller. The reason that some wavenumbers cannot be reached is related to the nature of the TDL output. This output consists of a number of modes with a fairly regular spacing and with varying amplitudes.

The number of modes is determined by the extent of the TDL gain curve. As the TDL temperature is tuned, these modes tune as much as a mode spacing, at which point the TDL cavity energy is sent into another mode. Certain operating points can be found where no modes are operating or where several modes are competing for the energy in an unstable fashion.

Once found, the operating points are a good guide as to how to obtain light at the desired wavenumbers during the same thermal cycle. These operating points do change with time and thermal cycling so that they may have to be optimized each time the TDL is used. The NO R(4.5) lines were accessible after the second thermal cycling, but the setting of the operating point was often tedious. A scan of these lines is shown in figure 3.

Operating Point History

The operating point history of available TDL's is important because it will allow us to plan those field deployment techniques which will make possible the easy transfer of a characterized TDL to a field TDLSS or the quick recovery of a specified wavenumber after a thermal cycling in the field. We discuss in some detail the history of the NO TDL. This TDL has been in use since March 1979 and has been kept cold continuously except for 5 thermal cycles. Most of these cycles were caused by power outages. We find in general that the operating point changes such that the heat sink temperature must be decreased and the drive current increased as a function of time to achieve the same wavenumber as before. The location of this wavenumber is taken to be the peak of the room water-vapor line at 1897.52 cm^{-1} . Figure 4 presents the history of this operating point. We found it convenient to present this data as a graph of how the TDL wavenumber would have changed if the operating point had not been changed. In order to calculate this hypothetical performance, it was necessary to determine the thermal and current tuning rates of this TDL.

The TDL current tuning rate was measured by scanning through the well known R(5.5) NO lines at 1897 cm^{-1} . The current tuning rate was $0.011 \text{ cm}^{-1}/\text{mA}$. The thermal tuning rate was calculated from the operating point data of figure 4 for the dates between 8/23 and 9/19. Since the same absorption line was used each time, the temperature changes must correspond to the current changes. Use of the current tuning rate leads to a thermal tuning rate of $1.2 \text{ cm}^{-1}/\text{K}$. The hypothetical performance of the TDL wavenumber was then calculated relative to the operating point x on 3/29, using these rates and the actual operating point changes for each date. This performance is shown in figure 4. Most of the changes were small, which indicated that the same mode was used for all measurements. This observation means that one could

recover operation on a given line by a simple tuning of the drive current. Changes during a cycle are typically less than 0.1 cm^{-1} . The large change in wavenumber after cycle 2 is probably an indication that the TDL characteristics change so as to modify drastically the mode structure. The magnitude of the change is probably about equal to the mode spacing in this particular TDL. In fact, after this thermal cycle, the TDL could be tuned so as to allow the measurements of both the R(5.5) and R(4.5) NO line transitions.

It may also be helpful for field deployment considerations to summarize the operating point optimization process after a thermal change. Figure 5 shows a composite view of this process after the second thermal cycle. Figure 5(a) notes the TDL output before cycling as measured by the grating spectrometer superposed on a TDLSS trace of the ambient water vapor lines. Figure 5(b) shows the grating spectrometer trace of the TDL output after the cycle at the same heat sink temperature, but at a drive current which places the output at the previous wavenumber. Figure 5(c) shows the grating spectrometer trace of the TDL output after both temperature and current have been optimized to give near single mode performance at the previous wavenumber. Optimization of the operating point after a thermal cycle may thus require optimization of the mode structure as well.

Wavenumber Nonlinearity of the TDL

The fine wavenumber tuning of the TDL is accomplished by the joule heating of the drive current. For small drive currents compared to the TDL bias current, the wavenumber tuning is nearly linear with drive current. Scans like figure 3 in general possess a nonlinear current drive axis which must be calibrated for accurate work. One cannot simply interpolate between NO lines to locate the N_2O line, for example, to the precision required (e.g. 10^{-4} cm^{-1}). The current drive axis is commonly calibrated through use of a scan of a stable etalon as shown in figures 6 and 7. The stable etalon used was a temperature stabilized Ge etalon commercially purchased with a finesse of 3 and a free spectral range of 0.048 cm^{-1} in the 1897 cm^{-1} region. Etalon fringe peaks are thus known to be equally spaced in the wavenumber domain.

One can expect a 5% nonlinearity over a 1 cm^{-1} tuning interval for 100 mA drive and 1 A bias.

We have studied the scan nonlinearities for a variety of purposes using our digital control and data acquisition system. Figure 6 shows an early scan of an etalon fringe pattern where the drive current at

each fringe peak has been recorded. The peak spacing in current steps and the tuning rate near each peak are listed in table II. This listing shows that the spacing increases as the drive current increases while the tuning rate decreases. These results are consistent with the joule heating picture. Figure 7 shows a digital record of a fringe scan overlaying the room water-vapor lines. We have analyzed this digital record with our etalon calibration software indicated in figure 2 and as outlined in table III. This more general analysis indicates that the tuning nonlinearity is more complicated than the quadratic expectation. We are continuing our analysis of the nonlinearities by trying to fit every point of the etalon curve rather than just the peak values.

Axial Mode Structure

Axial mode structure is important to the performance of a TDLSS for at least three reasons. First, the presence of several modes at the same operating point requires in general the use of a grating monochromator (or other mode selector) to select the desired mode. The use of a grating spectrometer serves to increase the complexity of the TDLSS, introduce transmission loss, and introduce optical noise, although it is convenient for making coarse wavenumber identifications. Second, the mode structure may be such that the mode selector may not be able to reject very near or very far modes. Third, the selected mode may have fine wavelength structure which can only be diagnosed by other techniques. The latter two situations would serve to diminish the specificity of the TDLSS.

We present here some representative results of spectral scans of the TDL output which are relevant to field deployment. Figure 8 shows a TDL mode scan which indicates the very wide range ($\sim 38 \text{ cm}^{-1}$) over which a mode selector may have to operate. Most TDL emission ranges span only about 10 to 15 cm^{-1} . Figure 9 shows a mode scan which indicates the closest spacing of modes that may occur (i.e. about 2.2 cm^{-1}). Potential modes are actually spaced by 1.1 cm^{-1} in this TDL, but operation at lower bias currents causes every other mode to drop out. At certain operating points, the selected single mode has fine structure. This fine structure has been studied in several ways, including heterodyne measurements with a local oscillator, calibration etalon scans, and scans of gas lines. In figure 10, we show examples of such scans of the NO doublet. The scan here happens to be the fm/second harmonic transform of the familiar NO doublet of $R_{11}(5.5)$. Figure 10(a) illustrates the mode structure (lower trace) and the absorption line scan (upper trace) when a single axial mode has been selected by the monochromator and when fine structure is not present. Figure 10(b) illustrates absorption line scans when the selected mode has fine structure. Each trace is taken with a different operating temperature and shows different fine structure. A typical spacing is 0.075 cm^{-1} or 2.25 GHz. Heterodyne measurements with a 10 micron TDL indicated

fine structure with separations of 250 MHz. (See figure 11). Scans of a stable etalon also provide information on the fine structure of TDL axial mode from the discontinuity of the fringe pattern. Figure 12 shows a scan of a solid Ge etalon with a finesse of 3 and a free spectral range of 0.047 cm^{-1} using a 940 cm^{-1} TDL. At the left, the trace has the expected amplitude and shape. However, in the center a discontinuity appears and the amplitude of modulation drops sharply. This suggests the TDL sends out several closely spaced modes at this point.

Figure 5a and figure 5c show typical monochromator scans of TDL output after operating point optimization. For a common slit setting (i.e. 0.27 mm or 1.6 cm^{-1}), we typically see about 5% of the TDL output in other distant modes. True single mode behavior to the 0.1% level was very seldom observed, as was the mode fine structure described above. Powers up to $20\mu\text{ W}$ were delivered to the detector in a single mode selected by the monochromator of 25% transmission.

DIAGNOSTICS OF TDLSS CHARACTERISTICS DUE TO TDL ENVIRONMENT

The characteristics due to TDL environment included short-term wavenumber stability, the instrument lineshape function, and intermediate-term wavenumber stability. The short-term wavenumber stability was measured by using the heterodyne technique which was described by G. Steinberg (ref. 1). This study was carried out using a TDL chosen to overlap the available CO_2 -laser lines near 10.6 microns. Figure 13 shows the heterodyne spectrometer and figure 14 shows the results achieved. The two lower traces indicate that the lineshape of the TDLSS is about 10 MHz in half-width (i.e. $\sim 3 \times 10^{-4}\text{ cm}^{-1}$) for averaging times up to several minutes. Single traces at the top left show that the central wavenumber remains constant to about 2 MHz. Single traces at the top right indicate that the wings of the lineshape are due to large wavenumber excursions caused by the residual vibrations from the closed-cycle cooler. These residual vibrations could not be eliminated. The goal of resolving Doppler lines was, however, substantially met. The lineshape study of Doppler and Voigt lines reported by Poultney et al. (ref. 3) confirmed the heterodyne measurements of instrumental lineshape. Figure 15 is the measured shape of a NH_3 line under Doppler conditions. The curve which connects the data points is the best fit Gaussian profile, with an instrument linewidth of $4 \times 10^{-4}\text{ cm}^{-1}$. The fitting procedure is described by Noll and Pires (ref. 4).

Intermediate-term wavenumber stability was estimated by using the TDLSS near 5.3 microns which was tuned to the half-width position

of NO $R_{11}(5.5)$, where the slope is the greatest. A sample cell of the spectrometer was filled with 0.9 Torr of NO, a pressure low enough for the absorption line to be Doppler limited, with a half-width at half maximum of about 0.002 cm^{-1} . The vertical deflection of the recorder pen then corresponds approximately to 0.004 cm^{-1} , or FWHM of the line. The diode current was then adjusted so that the diode wavelength was halfway down the absorption line; the point of maximum sensitivity to wavelength changes. Two traces are shown in figure 16. The upper trace shows the intermediate-term stability was about 0.001 cm^{-1} , with TDL temperature controller gain 50, while the short-term noise was much smaller. The intermediate fluctuation had a period of about 3 minutes. The short-term noise was larger ($\sim 0.0008 \text{ cm}^{-1}$), but the intermediate-term noise was smaller, with the TDL temperature controller gain 100.

WAVENUMBER CONTROL TECHNIQUE

The use of tunable diode lasers for laser heterodyne spectrometry requires that the issue of wavenumber control be addressed. There are various techniques to counteract the effects of thermal drift and thermal cycling of the laser. One of the most promising methods, based on counting and interpolating fringes of a stable Fabry-Perot interferometer, is described here.

The requirements for TDL wavenumber control imposed by spectroscopic and mission considerations have been summarized in table I. When the laser is turned on remotely, it must be tuned to the desired wavenumber even though its initial wavenumber may be uncertain due to thermal cycling. The tuning accuracy must be $\pm 10^{-4} \text{ cm}^{-1}$ to ensure operation at the center of a Doppler-broadened line. Ideally, a cell containing the gas to be detected would be available at all times to provide a reference absorption line at precisely the correct wavenumber. However, in many cases it is not possible to carry a sample of a particular species with a flight spectrometer. Such tuning accuracy is useful only if the technique can also reduce the intermediate- and long-term drift of the laser to less than 10^{-4} cm^{-1} .

Our concept for wavenumber control is depicted with the aid of figure 17. First the TDL wavenumber is swept across its tuning range until an absorption line in a reference gas cell is found. We assume that the TDL wavenumber remains within a mode tuning range of the reference line during thermal cycling and operation. The reference line need only be as close as 1 or 2 cm^{-1} to the desired line position. Then the TDL mode structure is optimized for operation in this spectral region. The TDL is locked to the gas reference line and the position (in wavenumber space) of its output with respect to the etalon transmission curve is recorded. Next the TDL is tuned across a predetermined number of Fabry-Perot interferometer transmission peaks to the wavenumber

of the spectral line of interest. Finally the TDL is stabilized at that wavenumber by monitoring transmission through the etalon, and long-term drift of the etalon is compensated by a separate control system.

Detection of reference gas absorption and interferometer fringes is accomplished by modulation of the TDL at a frequency f_1 , and modulation index A_1 . As shown in figure 18, the first and second fm/harmonic signals are available to identify the position in wavenumber space of the laser output. The precise location of the TDL wavenumber to within one hundredth of a free spectral range of the interferometer requires accurate interpolation between transmission peaks.

One simple, elegant method of interpolation uses information found in first and second fm/harmonic detection of the etalon transmission function (also known as the Airy function). For very low values of finesse the Airy function is nearly sinusoidal on top of a DC component, resulting in nearly sinusoidal fm/harmonics (see figure 19a). Since the first two fm/harmonics have a 90° relative phase difference, their ratio is approximately proportional to the tangent of a quantity which varies linearly with wavenumber σ , as shown in figure 19(b). Therefore, by computing θ , the arctangent of the ratio of the two harmonics of the Airy function, one obtains the wavenumber difference between the TDL output and any point on the etalon transmission function, except for an ambiguity due to the periodicity of the fringes. This ambiguity is resolved by counting the number of periods traversed as the TDL is tuned away from the reference line. Variations in TDL power do not present a problem because the ratio of fm/harmonic signals is used.

Use of moderate finesse etalons may degrade linearity below acceptable levels. The nonlinearity in the relationship between θ and σ is due to the deviation of the Airy function from a sinusoid. The nonlinearity can be reduced by judicious choice of the modulation index A_1 . As shown in figure 20, fm/harmonic detection is equivalent to filtering by a Bessel function in Fourier space. Here we use results of our complete analysis of the operation of a wavelength modulated spectrometer (ref. 5). The Fourier transform of the Airy function is a series of impulses at $0, 2L/c, 4L/c$, etc., of decreasing magnitude. The impulse at $2L/c$ represents the sinusoidal component of interest here; the others cause the nonlinearity. Detection of the first fm/harmonic, H_1 , results in filtering by $J_1(A_1)$; detection of the second fm/harmonic, H_2 , results in filtering by $J_2(A_1)$. Variation of A_1 greatly affects linearity, as is apparent from figure 21a. When A_1 is about 0.4 times the etalon free spectral range, a minimum in nonlinearity is achieved. This minimum is not sufficient to allow utilization of a moderate finesse etalon of practical length (e.g. 0.01 cm^{-1}).

Further improvement can be obtained by adding a second modulation at frequency f_2 , with modulation index A_2 . As shown in figure 20, the second modulation has the effect of filtering by $J_0(A_2)$, provided

the detection bandwidth is smaller than $|f_1 - f_2|$. The value of A_2 may be chosen to cancel the contribution of the delta function at $4L/c$, the major source of nonlinearity. The improvement in interpolation linearity is seen in figure 21b, for an A_2 of 0.19 times the free spectral range. Less than 1% nonlinearity may be achieved with reasonable etalon plate reflectivities of about 30%.

Drift of the etalon transmission curve, e.g. due to thermal expansion, can seriously degrade the accuracy of this technique. With a second laser (TDL 2), the drift of the etalon can be monitored and taken into account. If TDL 2 is continuously locked to a reference gas absorption line, its transmission through the etalon can accurately measure etalon drift. Use of fm/harmonic quadrature detection of TDL-2 transmission makes the monitor signal independent of TDL-2 power. The two lasers can be isolated by having their beams counterpropagating and orthogonally polarized. The detected etalon drift can provide continuous updating of the tuning of TDL 1, or can merely signal the need to repeat the TDL 1 wavenumber acquisition process.

In conclusion, a wavenumber control technique capable of meeting the most stringent requirements of a tunable laser heterodyne spectrometer system has been described and analyzed. Straightforward signal analysis principles have been used to adapt the method to practically realizable hardware. The requisite accuracy and long-term stability can be achieved with this system as long as the vernier etalon is stable between updates. Interpolation to 1% of a 0.01 cm^{-1} period etalon achieves the 10^{-4} cm^{-1} goal. The control technique and algorithms are discussed in greater detail by Wu and Poultney (ref. 6).

STABLE WAVENUMBER CONTROL ETALON

Our experience with solid Ge calibration etalons and adjustable air-space Fabry-Perot interferometers has led us to design a special purpose, air-space interferometer for the wavenumber control etalon. Table IV lists the features of this control etalon.

An air space etalon eliminates the refractive index inhomogeneity associated with thick germanium or silicon etalons. In addition, the temperature dependence of the index of air can be eliminated by sealing the space, and a spacer can be chosen with a lower coefficient of thermal expansion than that of a solid etalon.

Our air space interferometer uses a 25-cm reflector spacing to achieve a 0.02 cm^{-1} free spectral range. Figure 22 shows a schematic of the interferometer. Fabricating a 50-cm spacer would yield 0.01 cm^{-1} , which is our goal. The spacer is a fused silica tube, which provides inherently low thermal expansion. For fused silica, $\alpha = 5 \times 10^{-7}/^\circ\text{C}$. Only the so-called zero-expansion glasses have lower values,

and they are high cost items not readily available in tube form. Fused silica provides stability of $10^{-3} \text{ cm}^{-1}/^{\circ}\text{C}$.

The 2" outer diameter was chosen to be compatible with stock flat reflectors. The reflecting surfaces are optically contacted to the tube ends, assuring a moderately rugged, permanently aligned interferometer. The use of uncoated reflectors allows broadband performance. Zinc selenide provides 17% reflectivity and visible transmission for ease of system alignment. Germanium would give twice the reflectivity at the expense of this last feature, but we have determined the higher finesse unnecessary.

The outer surfaces of the reflectors are wedged with respect to the inner surfaces, and deliberate skewing of the wedges in both plates prevents the formation of a second etalon cavity. AR coatings could be utilized on the outer surfaces if experiment proves them necessary, but at the sacrifice of broadband performance.

With five-millimeter tube walls, the clear aperture is forty millimeters. Tube flex negligibly distorts the low finesse, even if the mounting is simple.

The design of the stable Fabry-Perot interferometer is described in more detail by McNally and Poultney (ref. 7).

CONCLUSIONS

We conclude that the characteristics of available laser diodes and aspects of their performance in closed-cycle cooler environments which may have made space flight systems impractical can be overcome through realizable control techniques. In particular, wavenumber purity and stability issues have been thoroughly diagnosed and a conceptual design has been developed for a wavenumber control system which can resolve performance questions related to these key issues.

REFERENCES

1. Steinberg, G.N.: Wavenumber Stability of a Laser Diode Mounted in a Closed Cycle Helium Refrigerator. *Rev. Sci. Instrum.*, vol. 50, no. 12, Dec. 1979, pp. 1622-1625.
2. Poultney, S.; Steinberg, G.; Pires, A.; Miller, M.; Chen, D.; Macoy, N.; Noll, R.; and Grossman, W.: Tunable Diode Laser Absorption Spectrometer as the Stratospheric Measurement System for the High Altitude Pollution Program - A Feasibility Study. Perkin-Elmer Report No. 14752, The Perkin-Elmer Corporation, Danbury, CT., 1980.
3. Poultney, S.; Steinberg, G.; Noll, R.; Pires, A.; Chen, D.; and Miller, M.: The Position of Water Vapor Lines near 5 Microns, NH Line Shape Measurements, and Comments on Wavenumber Calibration using Tunable Diode Laser Spectrometer. Published in 34th Annual Symposium on Molecular Spectroscopy, Ohio State University, Columbus, Ohio, June 11-15, 1979.
4. Noll, R.; and Pires, A.: A New Nonlinear Least Square Algorithm for Voigt Spectral Lines. *Applied Spectroscopy*, vol. 34, no. 3, May 1980, pp. 351-360.
5. Pires, A.; Poultney, S.; Chen, D.; and Miller, M.: Theory and Practice of the Wavelength Modulated Spectrometer. Preliminary report published in Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy, Atlantic City, New Jersey, March 10-14, 1980.
6. Wu, F.; and Poultney, S.: Wavenumber Control Techniques for Tunable Diode Laser Spectrometer. Perkin-Elmer Report No. 14762, The Perkin-Elmer Corporation, Danbury, CT., 1980.
7. McNally, M.; and Poultney, S.: Design of a Stable Fabry-Perot Etalon for TDL Spectrometer Calibration and Control. Perkin-Elmer Report No. 14763, The Perkin-Elmer Corporation, Danbury, CT., 1980.

TABLE I.- REQUIREMENTS FOR TDL WAVENUMBER CONTROL

- o Following Turn-On, TDL must Automatically Tune to Desired Wavenumber
- o Tuning Accuracy must be better than $\pm 10^{-4} \text{ cm}^{-1}$ (10% of Doppler Width), even if a reference cell with absorption line at desired wavenumber is not available
- o Short-term Drift must be less than 10^{-4} cm^{-1}
- o Intermediate-Term Drift must be less than 10^{-4} cm^{-1}
- o Long-term Drift must be less than 10^{-4} cm^{-1} for periods of hours
- o TDL Output at Detector must consist of Single Mode of Adequate Power

TABLE II.- PEAK SPACING IN CURRENT STEPS AND TUNING RATE
NEAR EACH PEAK

Fringe	TDL Current (Amp)	ΔI (ma)	$\frac{\partial \sigma}{\partial I}$ (cm ⁻¹ /ma)
1	1.7222	4.0+0.1	0.0120 = 0.0003
2	1.7262	3.9	0.0123
3	1.7301	3.9	0.0123
4	1.7340	4.1	0.0117
5	1.7381	3.9	0.0123
6	1.7420	4.1	0.0117
7	1.7461	4.1	0.0117
8	1.7502	4.2	0.0114
9	1.7544	4.2	0.0114
10	1.7586	4.2	0.0114
11	1.7628	4.2	0.0114
12	1.7670	4.3	0.0112
13	1.7713	4.3	0.0112
14	1.7756		

TABLE III.- PROCEDURE FOR WAVENUMBER SCALE CALIBRATION

1. Find Etalon Peak Locations

- o locate peak vicinity
- o quadratic fit fringe to locate fringe center

2. Expand peak positions in nth order polynomial

- o assign fringe number to peak
- o least-squares fit nth order polynomial

$$y = a_0 + a_1x + a_2x^2 + \dots$$

where x = fringe position

y = fractional fringe number

3. Calibrate wavenumber scale using known FSR and reference line

- o $\sigma(x) = \text{FSR} \cdot [a_0 + a_1x + a_2x^2 \dots] + \sigma_{\text{ref}}$

TABLE IV.- ETALON DESIGN FEATURES

- o Inherent Temperature Stability — $10^{-3} \text{ cm}^{-1}/^{\circ}\text{C}$ with fused silica spacer
- o Short Free Spectral Range — 0.02 cm^{-1} with 25 cm cavity
- o Permanent Alignment — Solid Assembly of Reflectors to Spacer
- o Simple System Insertion — Transparent in Visible for Alignment
- o Broadband Reflectivity — Available through Fresnel Reflection
- o Large Aperture — 40 mm C.A.
- o Freedom from Secondary Etalon Effects — Uses Wedged Reflectors
- o Freedom from Index Inhomogeneity — Assured by Closed Air-Space Design

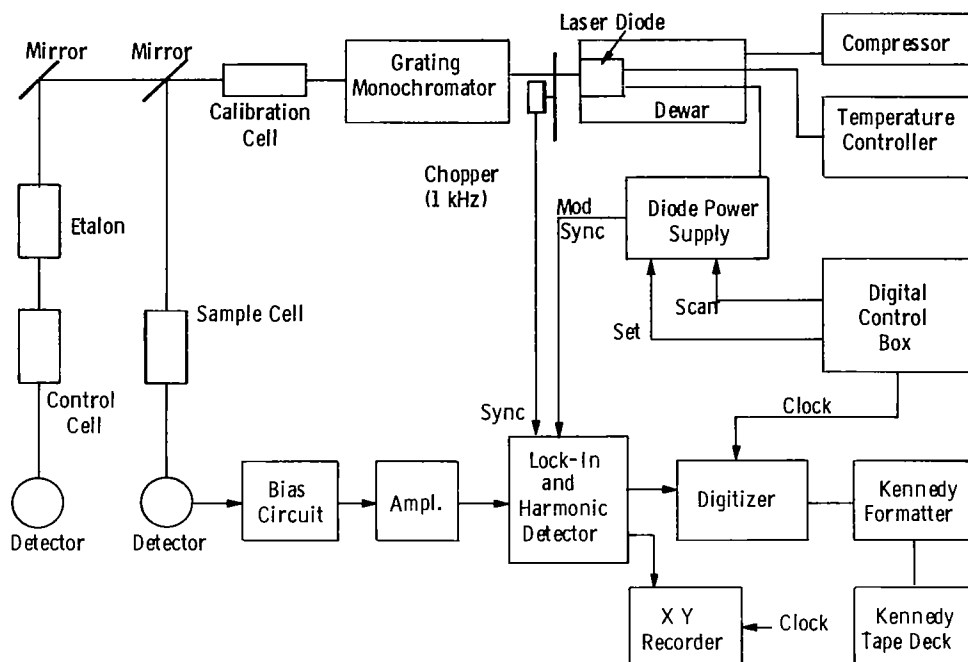


Figure 1.- Tunable diode laser absorption spectrometer; uses all reflective optics and can provide graphical or IBM-compatible digital records.

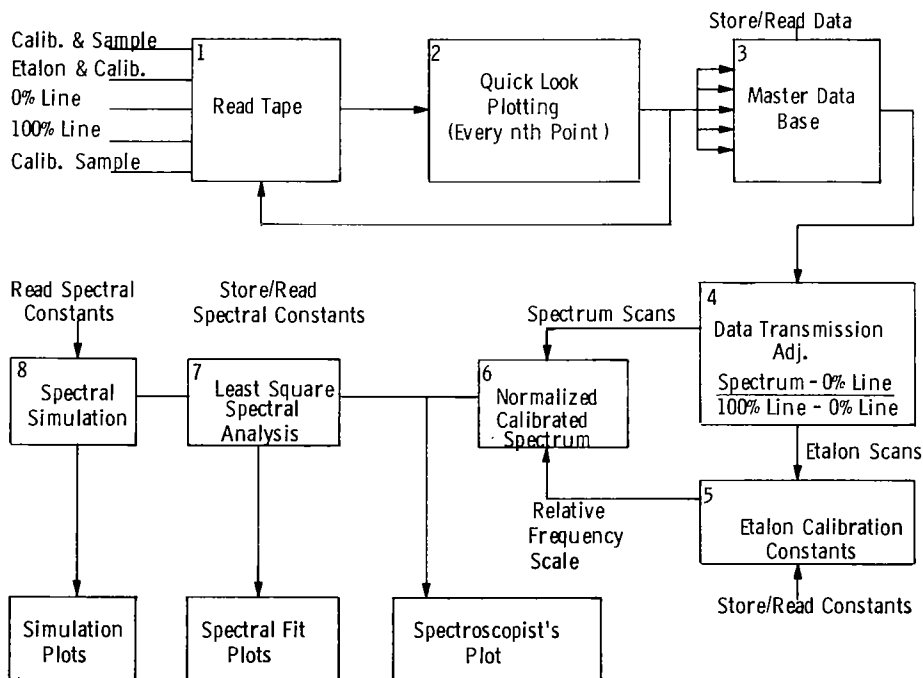


Figure 2.- Flowchart of executive spectroscopy software for tunable diode laser absorption spectrometer.

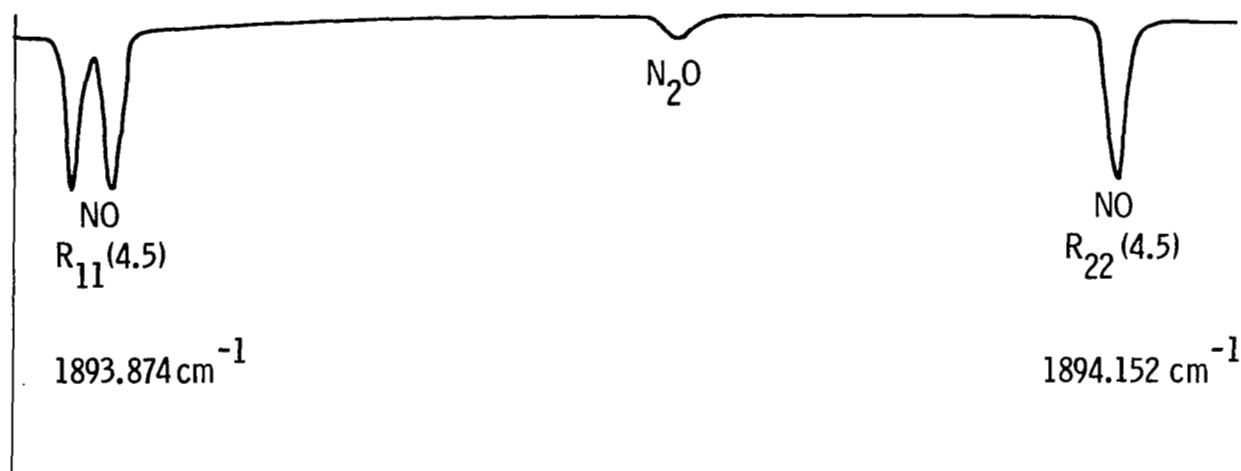


Figure 3.- Laboratory NO absorption spectrum for $R(4.5)$ transition near 1894 cm^{-1} . Also shown is weak N_2O line due to 10 Torr N_2O added to second sample cell.

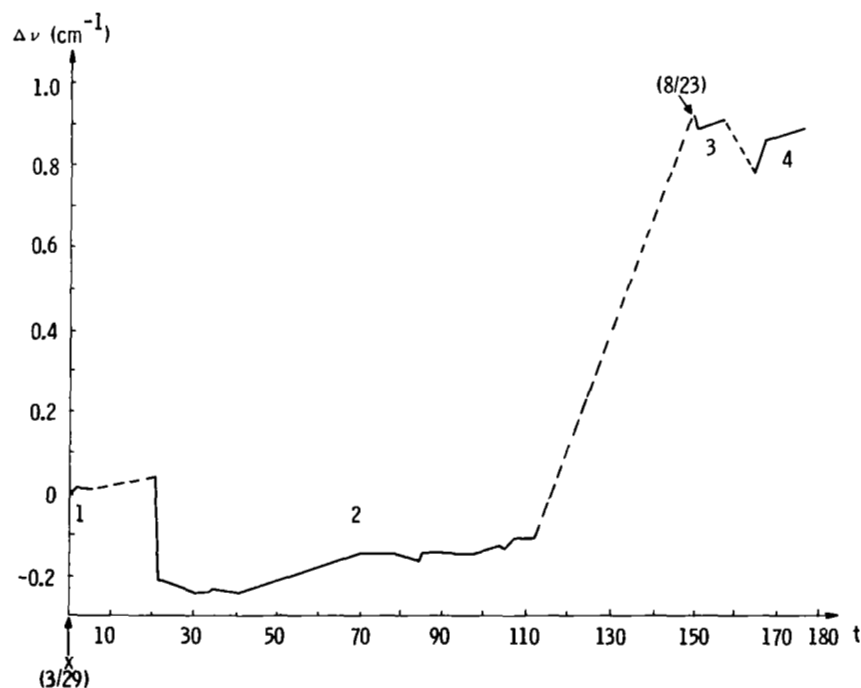


Figure 4.- TDL operating point history. During operation, operating point would change by about 0.1 cm^{-1} over longer term, with even larger changes occurring during some thermal cycles.

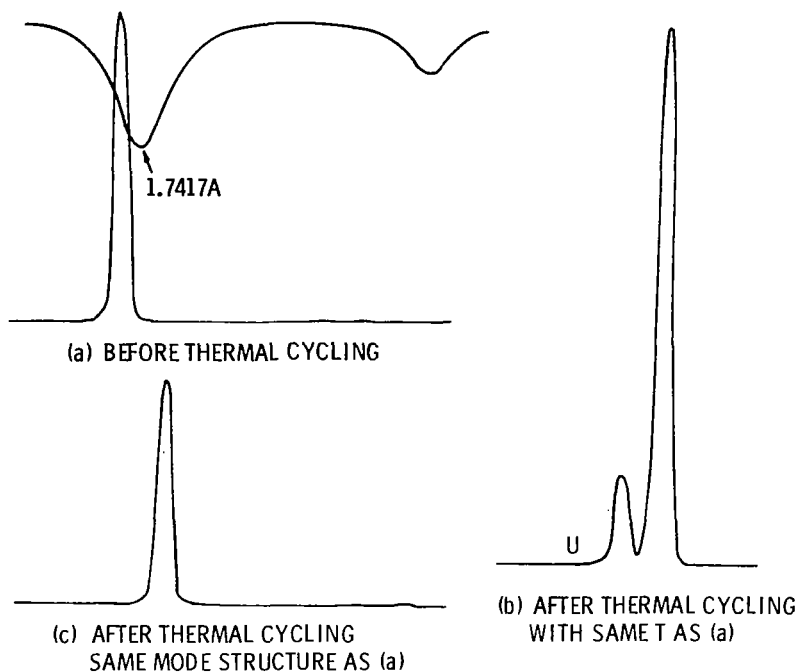


Figure 5.- Optimization of TDL mode structure after thermal cycling. After each thermal cycle, TDL operating point changes such that heat sink temperature must be decreased and drive current increased to achieve same wavenumber and to obtain as much power in single mode as possible.

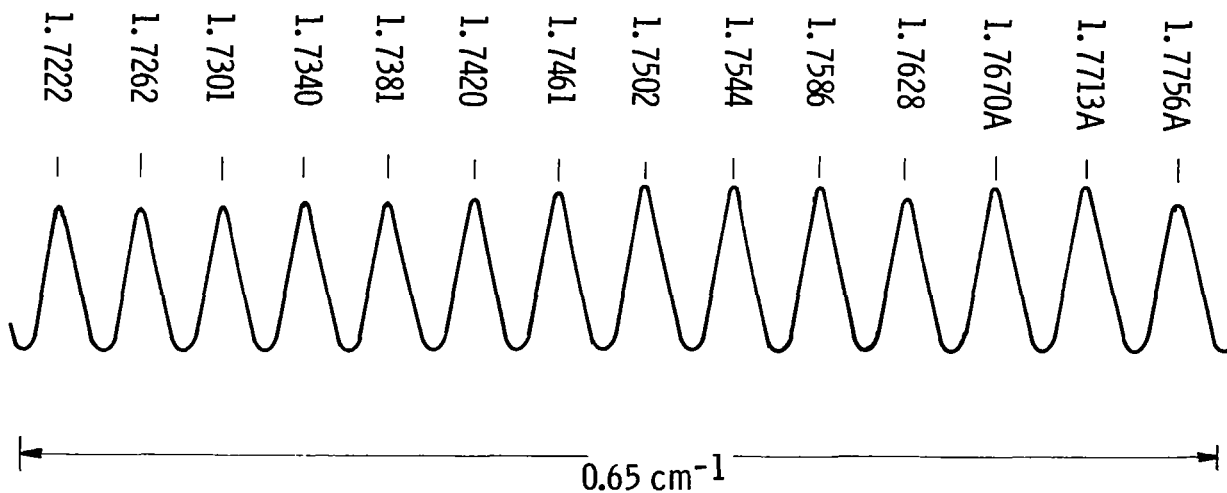


Figure 6.- Germanium etalon fringe scan. Drive current at each fringe peak is indicated. FSR of etalon is 0.048 cm⁻¹.

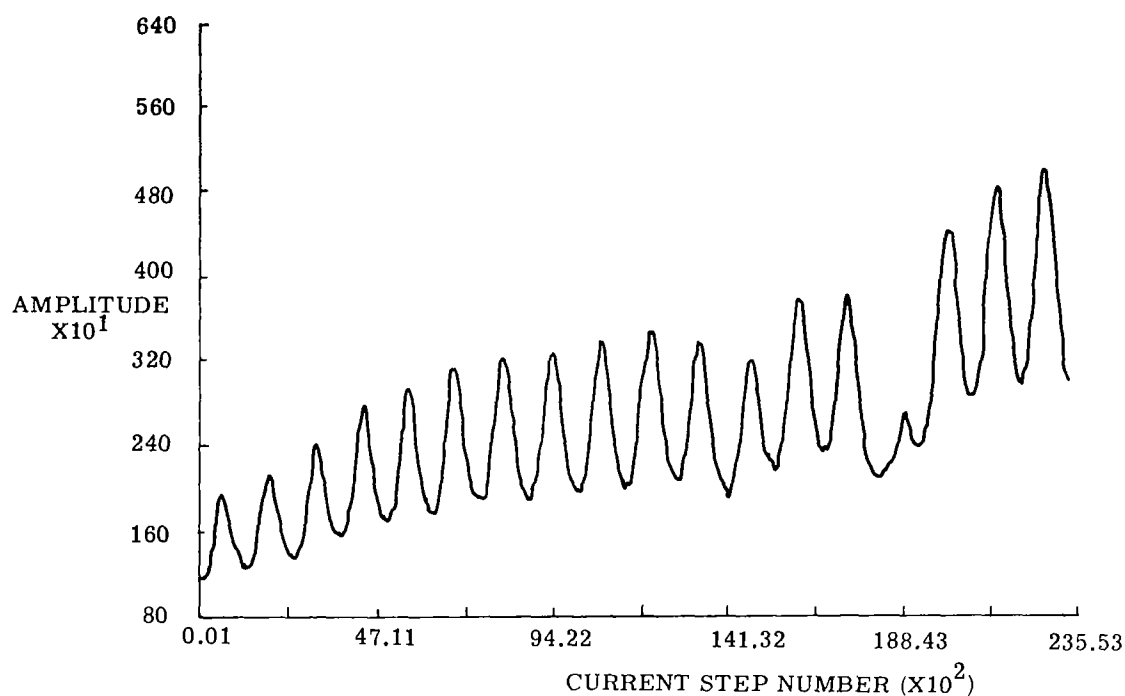


Figure 7.- Digitized etalon scan. Etalon fringes distorted by H_2O lines are discarded from fitting process.

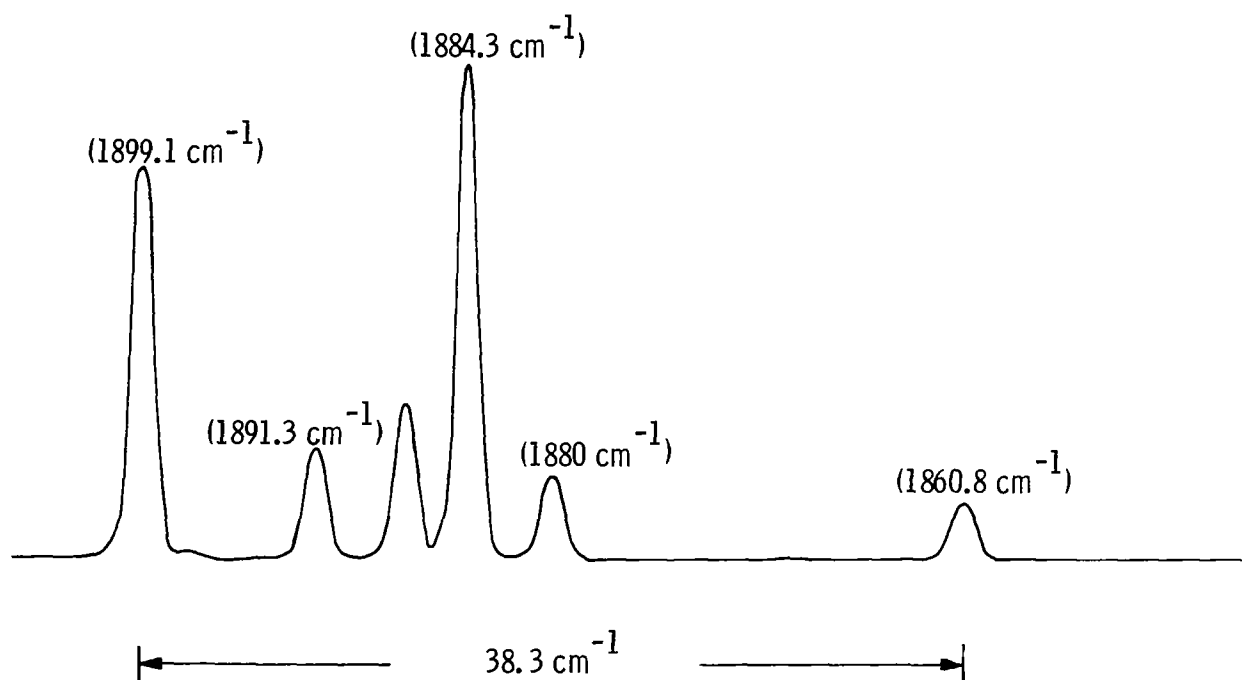


Figure 8.- TDL mode scan; indicates very wide range over which mode selector may have to operate.

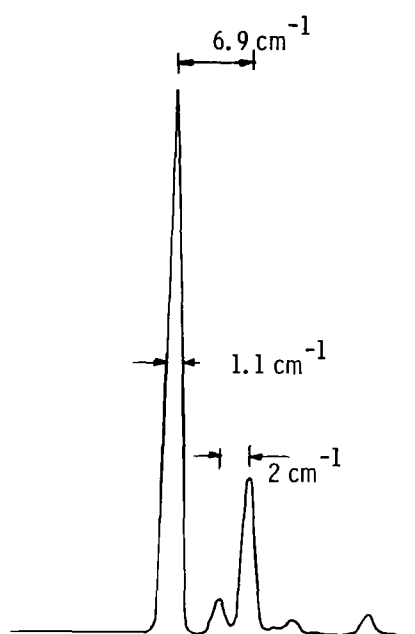
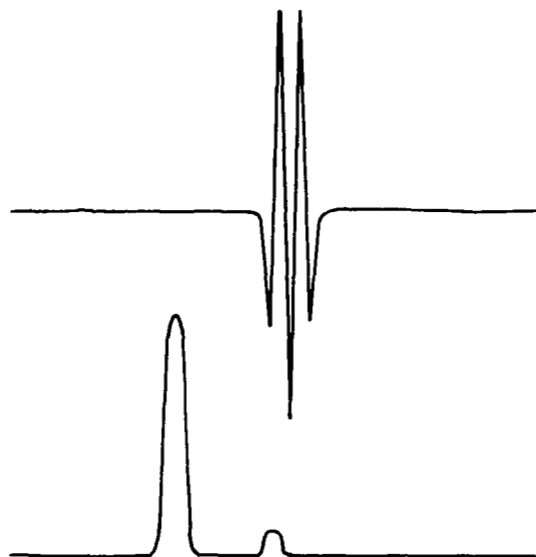
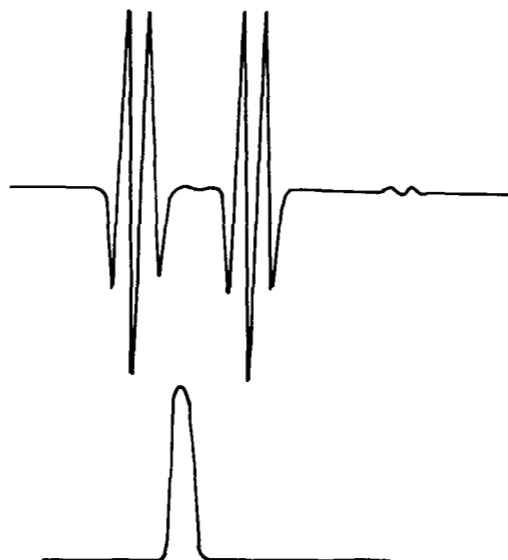


Figure 9.- TDL mode scan; indicates closest spacing of modes which may occur (i.e. about 2 cm^{-1}).



(a) Single axial mode selected by monochromator when fine structure is not present.



(b) Absorption line scans when selected mode has fine structure.

Figure 10.- Absorption line scan evidence for fine mode structure, showing mode structure (lower trace) and absorption line scan (upper trace). Each trace is taken with different operating temperature and shows different fine structure. Typical spacing is 0.075 cm^{-1} or 2.25 GHz .

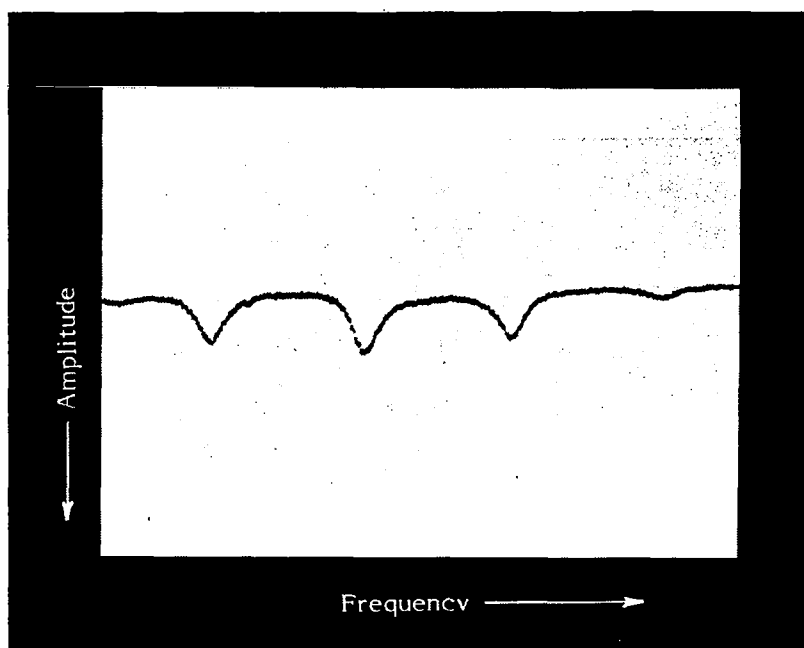


Figure 11.- Heterodyne beat between CO₂ laser and diode laser fixed near 10.6 micrometers showing fine structure of single axial mode. Peaks are separated by 250 MHz.

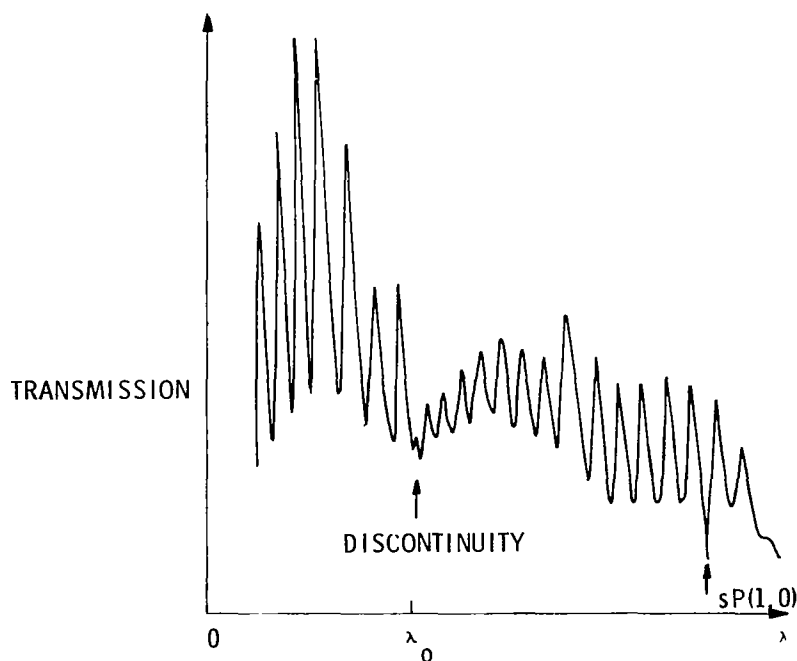


Figure 12.- Calibration etalon evidence for mode fine structure with 10.6 microns TDL. Note discontinuity at marked λ_0 .

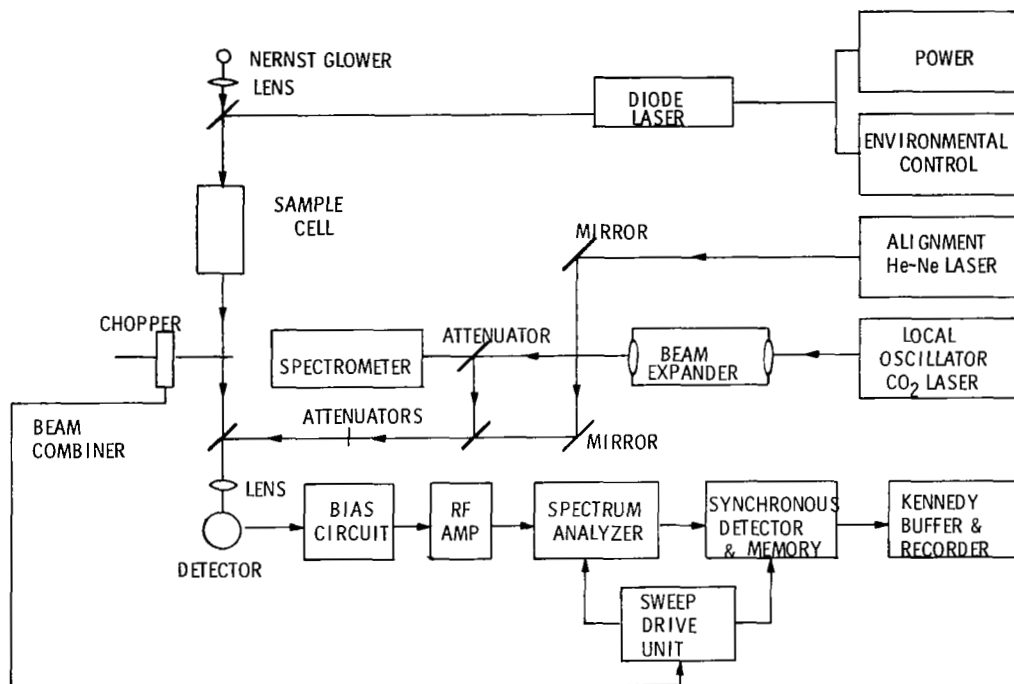


Figure 13.- Schematic diagram of heterodyne spectrometer with digital synchronous detector.

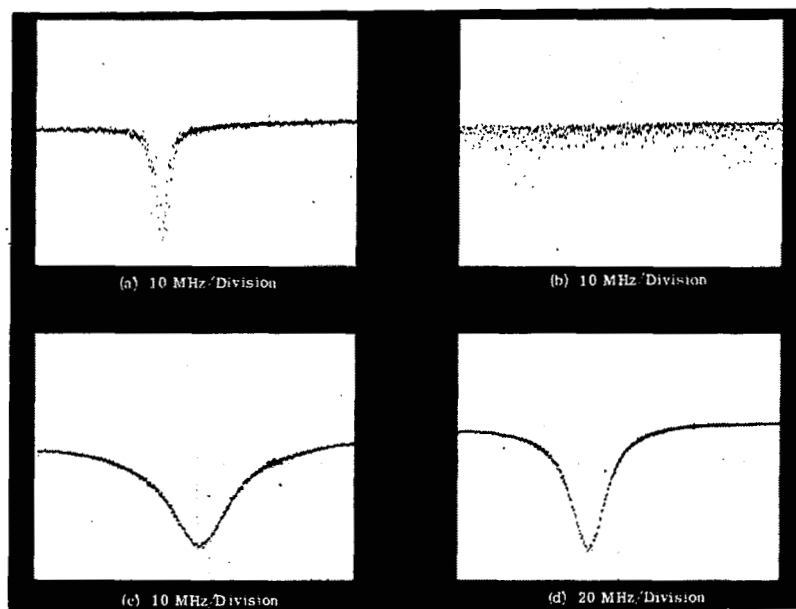


Figure 14.- Short-term stability of laser diode output wavelength. Results obtained with CO₂ laser heterodyne spectrometer for time period stated. All frequency sweeps are 10 MHz/div. except for (d) where it is 20 MHz/div. Equivalent line shape is shown to be 20 MHz FWHM in (c) and (d), for 75-s averaging time and is thought to be due to residual vibrations. Both (a) and (b) are selected single sweeps.

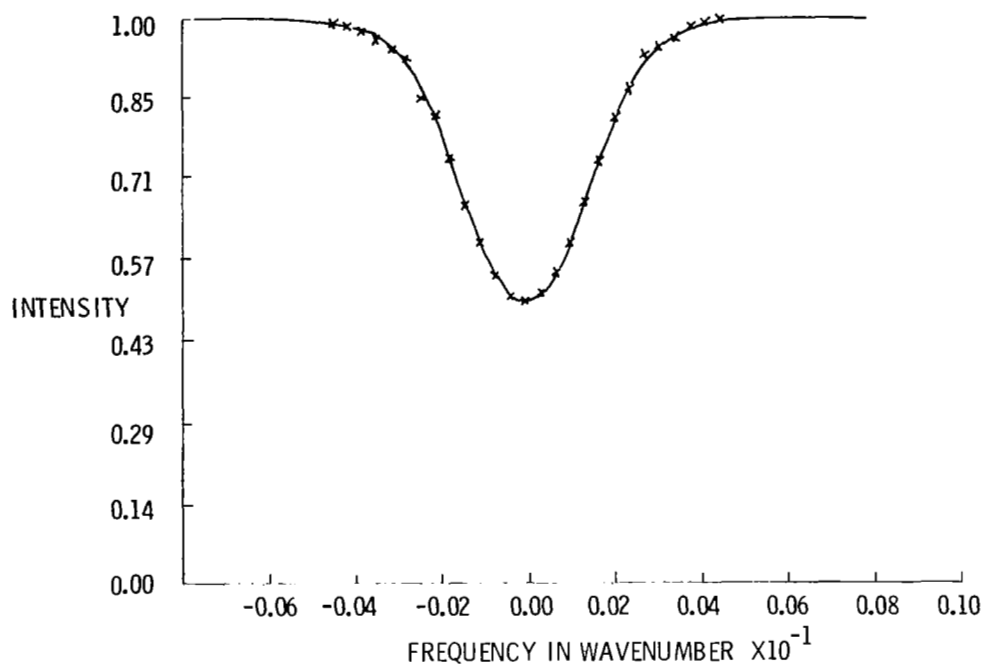


Figure 15.- Shape of $sP(1,0)$ line of NH_3 near 948 cm^{-1} under Doppler conditions in 2.5-cm cell. Points indicate data. Curve is best fit using least-square technique.

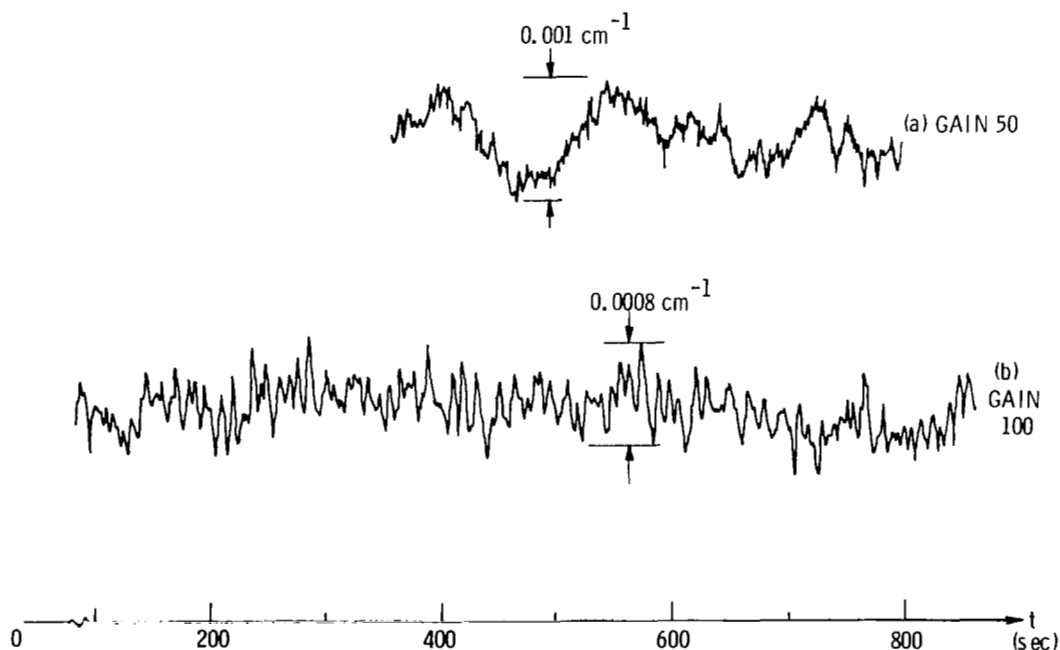


Figure 16.- Intermediate-term wavenumber stability; about 0.001 cm^{-1} with TDL temperature controller gain 50 (upper trace). Lower trace shows stability 0.0008 cm with TDL temperature controller gain 100.

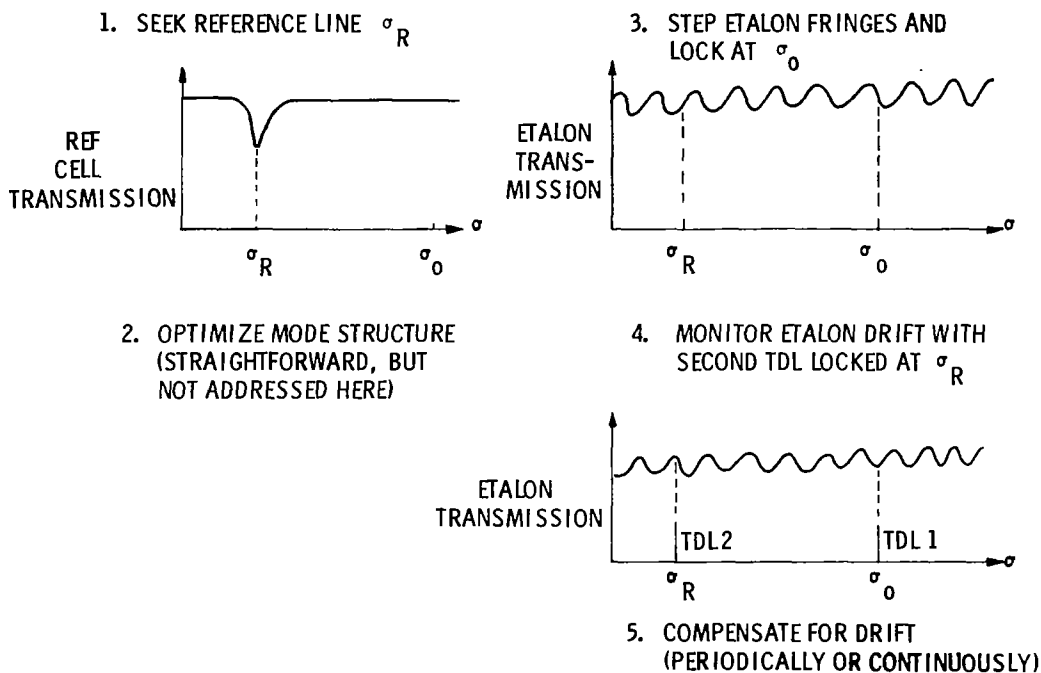


Figure 17.- Use of stable etalon for wavenumber control. σ_R is wavenumber of reference gas absorption line; σ_0 is desired operating wavenumber.

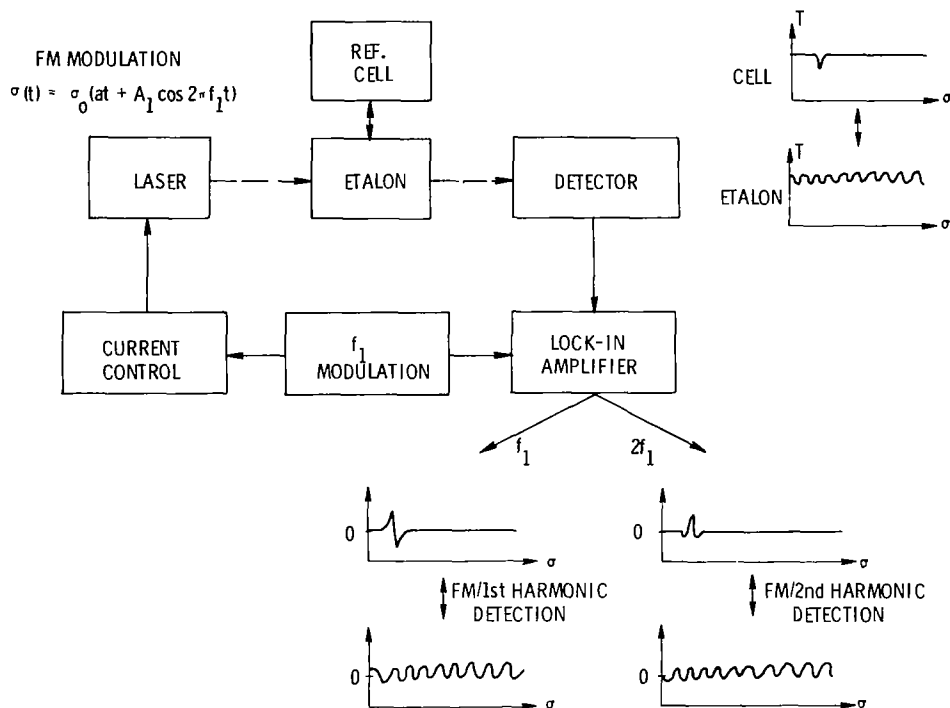
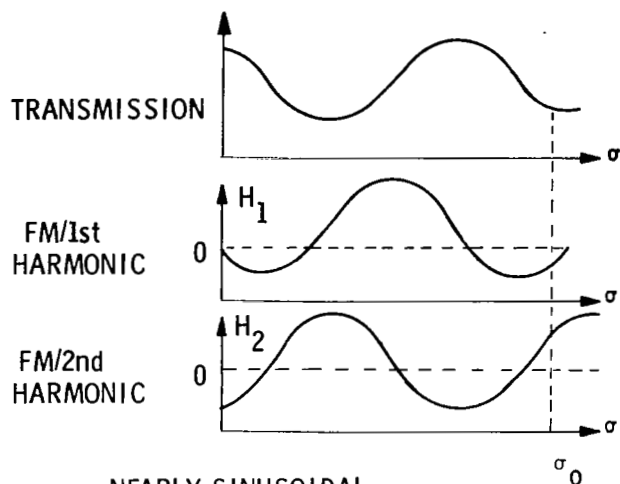
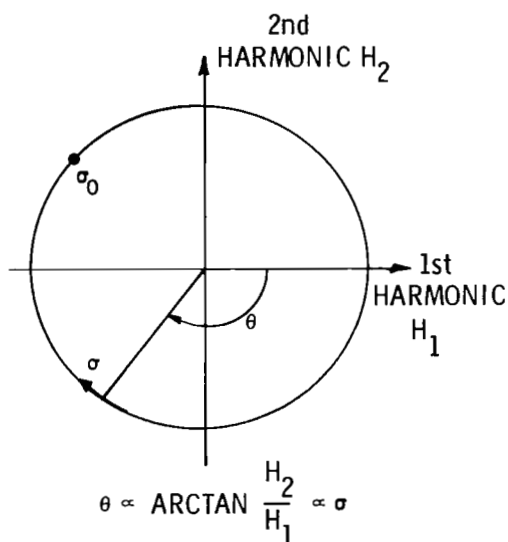


Figure 18.- Schematic of wavenumber control system. Transmission of reference cell is monitored until TDL is locked to reference line. Then etalon transmission is monitored for fringe position and stepping.



- NEARLY SINUSOIDAL
- 1st AND 2nd HARMONICS ARE $\frac{\pi}{2}$ OUT OF PHASE
- θ IS INDEPENDENT OF TDL POWER

(a) Etalon transmission function and its fm/harmonic detection signals H_1 and H_2 .



(b) Relation between H_1 , H_2 , and θ , which is proportional to wavenumber (approximately). As laser is tuned, θ increases linearly with wavenumber.

Figure 19.- Etalon fringe interpolation.

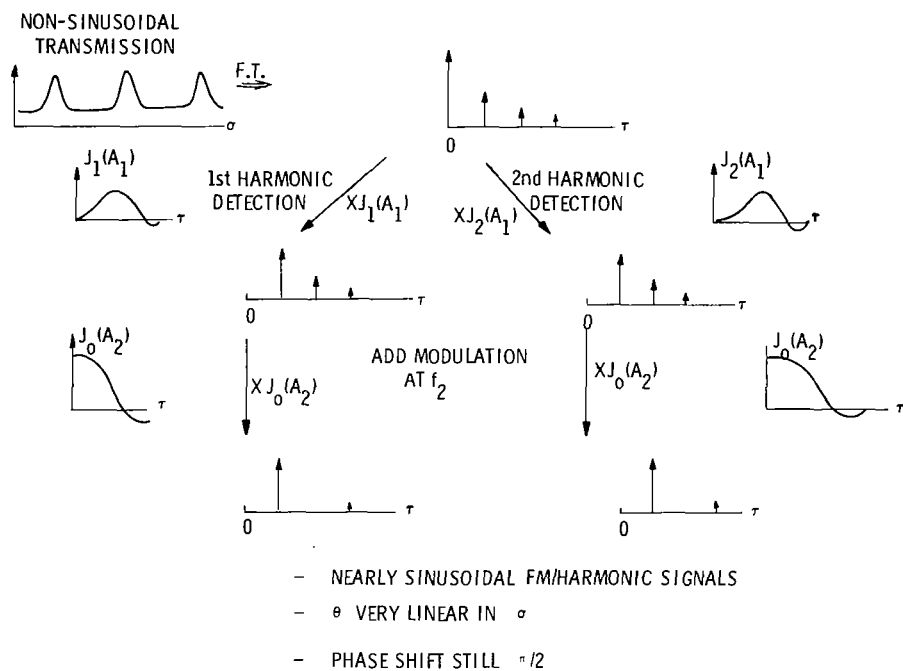
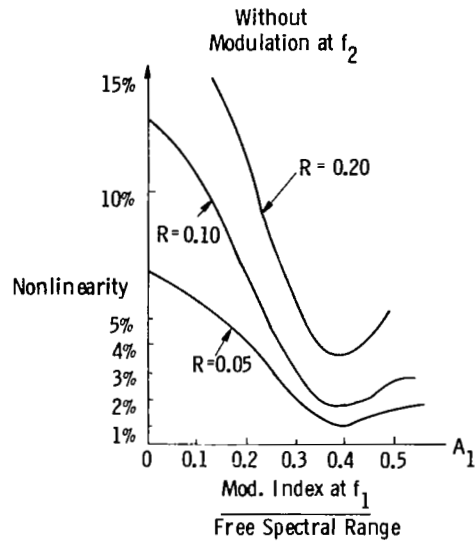
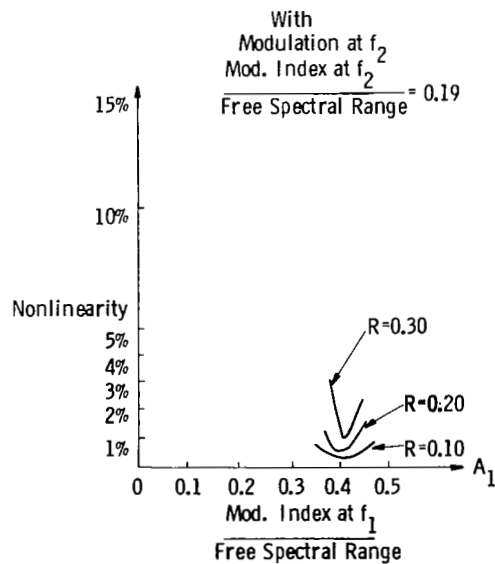


Figure 20.- Filtering by Bessel functions for improved linearity. TDL is modulated with modulation index A_1 at frequency f_1 and index A_2 at f_2 . Demodulation at first harmonic (f_1) and second harmonic ($2f_1$) corresponds to filtering by $J_1(A_1)$ and $J_2(A_1)$, respectively, followed by $J_0(A_2)$.

$$\text{Nonlinearity} = \frac{\text{Max. Error of Linear Approximation}}{\text{Free Spectral Range}}$$

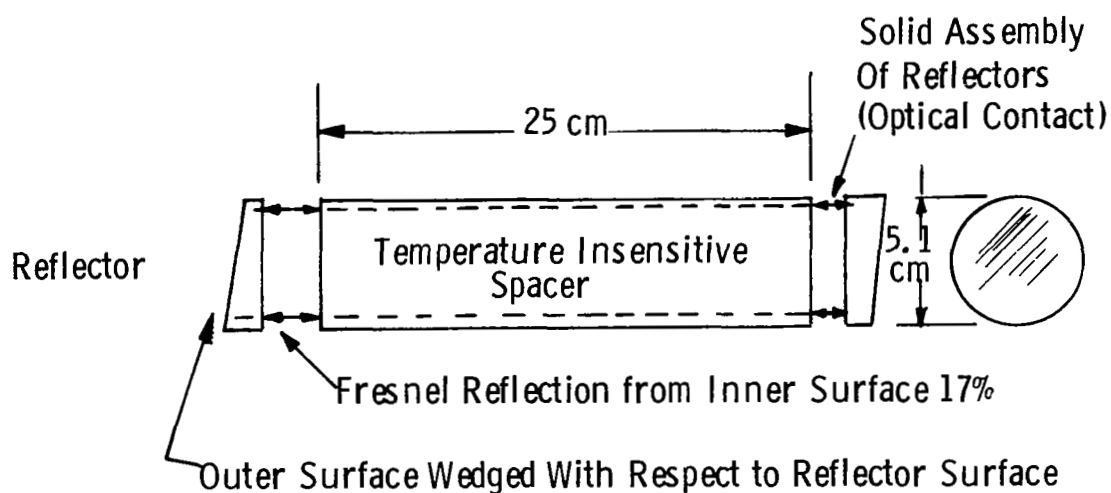


(a) Effect of modulation index A_1 at frequency f_1 on interpolation linearity.



(b) Same as (a), but with second modulation at f_2 . Much better linearity can be achieved.

Figure 21.- Interpolation linearity.



Tube Flexing has Negligible Effect on Finesse - - Mounting is Non-critical

Figure 22.- Air space etalon; constructed of two zinc selenide flats, contacted to fused silica tube which acts as spacer. Fresnel reflection from inner surfaces of ZnSe plates is used, and outer surfaces are wedged to eliminate accidental etalon behavior.

EXCESS NOISE IN $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ SEMICONDUCTOR LASERS

Charles N. Harward
Department of Physics
Old Dominion University

and

Barry D. Sidney
NASA Langley Research Center

INTRODUCTION

Tunable Diode Lasers (TDL) have been used for several years as ultra narrow band (<30 MHz) sources for ultra high resolution spectroscopy of many different infrared absorbing gases. Interest in the TDL has been increasing in recent years due to their possible use in different types of spectroscopic systems, i.e., Laser Absorption Spectrometer (LAS) and Laser Heterodyne Spectrometer (LHS). These systems all have the common goal of making remote or in situ measurements of one or several atmospheric gases and use the TDL as a spectroscopic source. Because of the interest in high detection sensitivities, the noise characteristics of the TDL have been studied for frequencies less than 20 kHz (refs. 1 to 3). For heterodyne applications, the high frequency (>1 MHz) characteristics are also important. Therefore, we have studied the high frequency noise characteristics of the TDL as a part of a full TDL characterization program which has been implemented at Langley Research Center for the improvement of the TDL as a local oscillator in the LHS system. In this study which involved the characterization of TDL's from two commercial sources (Laser Analytics and New England Research Center) and one private source (General Motors Research Laboratory), it has been observed that all the devices showed similar high frequency noise characteristics even though they were all constructed using different techniques. We will now report these common high frequency noise characteristics.

EXPERIMENTAL DETAILS

The TDL noise characteristics facility is shown in figure 1. The TDL is mounted in a commercial closed cycle cooler which has been modified to minimize frequency fluctuations which are induced during the cooling cycle. Radiation is coupled through a wedged AR coated ZnSe window to a 37 mm f/1 Ge lens which collimates the TDL emission. This collimated beam passes either to a high speed HgCdTe detector (IF bandwidth ≈ 2 GHz) or to a monochromator. The monochromator was used to investigate the spectral characteristics of the TDL output. The high speed detector was used to measure the high frequency noise characteristics of the TDL output. The detector high frequency output was amplified by two cascaded IF amplifiers which were followed by a square law detector used to

rectify and integrate the IF output. The low frequency portion of the detector output and the integrated IF were monitored on a dual channel oscilloscope. At times, the frequency distribution of the IF was displayed on a spectrum analyzer.

The high frequency noise present in the TDL output for the numerous devices studied can be categorized according to frequency characteristics. Figure 2(a) shows a typical example of what we have categorized as low frequency noise (<100 MHz). The spectrum analyzer was set on 10 MHz/div with a 300 kHz resolution bandwidth and 10 dB/div. The zero frequency is located at the start of the left to right scan. For this particular spectrum, the noise drops to the constant baseline value in the first 30 MHz. In other cases, the noise extended out to 100 MHz. This type of noise can be caused either by a small amount of optical feedback or by physical phenomena intrinsic to the semiconductor laser itself. In the example represented in figure 2(a) the noise is associated with low frequency switching between laser modes (or sets of modes) at a particular current and temperature.

The second category of noise is the type which contains one or several narrow band spikes. In the case of several spikes, the spikes sometimes form a harmonic sequence (ref. 4). This type of structure is shown in figure 2(b). The total frequency was 1.2 GHz. The noise shown here is sometimes associated with the frequency of intensity self-pulsations (refs. 5 to 8) and is called self-pulsation noise. The harmonics appearing in the noise spectra may come from the waveform distortion in the pulsing light output (ref. 9). This category of noise is also associated with a heterodyne beating between closely spaced transverse modes. Harward and Hoell (ref. 10) have reported this type of harmonic structure for Pb-salt lasers with optical feedback while Broom and others (refs. 11 to 14) have shown this to be true for GaAs semiconductor lasers. Figure 2(c) shows the harmonic structure induced by optical feedback. The spacing between the peaks is commensurate with the spacing between the feedback element and the TDL.

The case of a single narrow frequency spike is shown in figure 2(d). This type of structure has been the subject of experimental and theoretical investigations in GaAs (refs. 15 to 19). The theoretical treatment of McCumber (ref. 15) for fluctuations in multimode laser oscillators was based on simple rate equations. The intrinsic quantum fluctuations drive the coupled electron and photon distributions at a characteristic resonance spiking frequency. This type of noise is called resonant shot noise.

The last category of noise structure has a broadband nature. This is shown in figure 2(e) for a 0 to 1.2 GHz spectrum. The noise structure is seen to rise above the no-noise baseline (bottom trace) in a more or less uniform manner. The broadband structure may be associated with the "low" frequency tail of a resonance where the spiking frequency is greater than the maximum frequency of the spectrum analyzer. This type of tail has also been seen in GaAs (refs. 13 and 20).

The next figures show the spatial dependence of both the direct detected power and the IF integrated noise. The bottom and top trace in each of the photographs in figures 3(a), (b), and (c) are, respectively, the TDL direct

detected power and the IF integrated noise present in the TDL output. For these measurements an etalon is inserted into the beam path and scanned through one free spectral range. Each of the figures 3(a), (b), and (c) are for different spatial positions in the TDL image. Each spatial position produces a different modal distribution of both the TDL power and the IF noise. The differential translation between each measurement was about 25 μm . The highest power TDL mode does not have the highest IF noise. In fact, numerous cases have been observed in which a very small mode has a very large amount of noise. Each of the figures 3(a), (b), and (c) were taken at the same temperature and current conditions on the TDL. The observed spatial dependence of the mode and noise mode structure has been observed in GaAs and has been attributed to different lasing filaments in the laser junction (refs. 18 and 21). Due to the aberrations in the present optical system, the individual lasing filaments cannot be optically separated. Their presence is substantiated by the spatial dependence of the TDL modes.

Figure 3(d) shows the output for the same spatial position as in figure 3(c), but with the etalon removed. All TDL modes present in figure 3(c) are present on the detector simultaneously in figure 3(d). The bottom trace, which was AC coupled, shows a variation not observable in the previous cases because the total power level is much higher. The variation was caused by mechanical vibrations in the closed cycle cooler. The upper trace, which was DC coupled, shows that the noise level is much reduced (in this case to zero) when all of the TDL modes are incident on the detector. The figure shows that the high frequency noise of a single mode which is caused by high frequency intensity fluctuations, is, for a multimode laser, higher than fluctuations of the total output. This result occurs because the strong negative correlations between the single, coherently oscillating mode and the nonlasing modes act to reduce the total noise output. This result, too, has been observed in GaAs semiconductor lasers (refs. 20 and 22). Both of these references have reported single mode noise power levels 30 dB higher than the noise power for all modes. The noise in this case has been given the name partition noise and is associated with multimode operation. The decrease in noise when multimodes are incident on a detector compared to the noise in a single mode is called noise amplitude stabilization.

In the previous set of figures, the spatial dependence of TDL mode structure was demonstrated along with the phenomena of noise amplitude stabilization. In figure 4, parts (a), (c), and (e) also show the spatial dependence of the mode structure and the IF noise while parts (b), (d), and (f) show the spatial dependence of the noise amplitude stabilization. Figures 4(a), (c), and (e) were obtained in the same manner as figures 3(a), (b), and (c) (for a different current and temperature). The variation in the individual mode power with spatial position as well as variation in the IF integrated noise is again observed. Figures 4(b), (d), and (f) were obtained by synchronizing the oscilloscope to a mechanical chopper which replaced the scanning etalon at each position. (The chopper was used to determine the noise baseline with no TDL radiation on the detector.) These figures were obtained at the same positions as figures 4(a), (c), and (e), respectively. Figure 4(b) shows only a slight amount of noise when all modes are incident on the detector even though the individual modes appear to be much noisier. At this position, the noise is again amplitude stabilized as in figure 3(d). As the detector is translated

across the TDL image to obtain figures 4(d) and (f), the amount of IF noise variation is much larger than the variation in the direct detected power, indicating a spatial dependence of the noise amplitude stabilization.

Figures 5(a), (b), and (c) show the spatial dependence of the TDL modes and the IF noise in these modes for a different current and temperature than was used in the previous figure. These figures were also obtained by moving the high speed detector to different positions in the image of the TDL. The difference in the TDL modes and the IF noise generated by these modes is evident for the three spatial positions. These results are similar to those shown in figures 3(a), (b), and (c), and 4(a), (b), and (c), except that two sets of longitudinal modes have been identified in the direct detected TDL mode spectra. These are indicated by the two sets of equally spaced lines at the top of each figure. The breaking up of the TDL output into two groups of longitudinal modes is suggested by the appearance of the IF integrated noise in figure 4(a) which is dominated by the noise generated by one set of longitudinal modes. The appearance of two sets of longitudinal modes is a common occurrence in GaAs lasers and has been attributed to different lasing filaments (refs. 21, 23, and 24). Each filament may act as an independent laser with its own family of modes. The different characteristic wavelengths are due to material inhomogeneities. Optical coupling between filaments was reported by Deutsch and Hatz (ref. 25) as well as uncoupled filaments by Guekos and Strutt (ref. 21).

Figures 6 and 7 indicate that the interaction between lasing filaments is involved in the appearance of excess noise in the laser modes. The curves in figures 6(a) to (e) show monochromator spectral mode scans of the TDL output from 9.35 to 9.47 μm for different TDL injection currents. The current in scan (a) was 1.337 amps and in each succeeding scan, the current was increased by 1.0 mA. Scans (a) to (c) show little variation in the mode structure. The TDL output was made up mainly of the modes labeled 1, 4, 5, and 7. At $I = 1.340$ amps, modes 2, 3, and 6 appear. The relative spacing of all the modes indicate modes 2, 3, and 6 belong to one lasing filament while 1, 4, 5, and 7 belong to another. Table I shows how the power in the individual modes changes with current. The total power in the modes remains relatively constant, with the power decrease in modes 1 + 4 + 5 + 7 matched by an increase in modes 2 + 3 + 6. The large change in power in the modes occurs at $I = 1.340$ amps where modes 2, 3, and 6 appear. The appearance of these modes coincides with the occurrence of broadband excess noise in all the TDL modes. Thus, the noise is associated with the appearance of the second filament and is probably due to an interaction between the filaments either in the TDL, caused by coupling of the filaments, or on the detector, caused by a beating between the filaments. The first case is the more likely since the noise was observed to be broadband rather than of the spiking resonance type, and the relative separation of the individual modes gives rise to beat frequencies which are higher than our system bandwidth.

Figure 6 shows that the appearance of a new filament, as indicated by the appearance of new laser modes, coincided with the onset of excess noise, thus implying that the onset of filamentary operation was responsible for excess noise. The next figure shows that filamentary operation, per se, does not necessarily give rise to excess noise. Figure 7 shows the same type mode scan as in figure 6. In figures (a) to (l), the current is changed from 1.7895 to

1.8005 amps in 1 mA steps. The spacing between modes indicates modes 1, 3, and 5 belong to one set of axial modes associated with one filament while modes 2, 4, and 6 belong to another set and, thus, another filament. The excess noise is not present in the TDL output in mode scans (a) to (e). Starting with scan (f) and continuing, the TDL output shows excess noise. Examination of table II shows that while changes occur in the individual modes, the relative power in each set of modes, as well as the total power, remains relatively constant. The exact cause of the noise is not known. One possible explanation is that, initially, the two filaments were sufficiently spatially separated so that they were uncoupled and noninteracting. The current increase from 1.7935 to 1.7945 amps may then have caused a slight spatial shift of the filaments sufficient to cause interaction between them with the consequent production of excess noise. Both uncoupled and coupled filaments have been observed in GaAs (refs. 21 and 25).

Figure 8(a) shows a TDL operating multimode with no excess noise in individual modes. The bottom trace of the figure shows the direct detected TDL mode structure obtained with the scanning etalon. The top trace shows no excess noise in the integrated IF for the individual TDL modes. Moving the detector about in the TDL image did not show a change in the mode structure as has been observed in previous figures, indicating only one lasing filament. Thus, this is one more indication that excess noise may in some cases be caused by an interaction between lasing filaments. Figure 8(b) shows the output of the TDL which is also quiet when all the modes are incident on the detector. In this figure, the scanning etalon was replaced by a mechanical chopper as in some of the previous figures. Since the individual modes are quiet, any one of these modes may be optically isolated and used as a spectroscopic source in a high speed detection system.

CONCLUDING REMARKS

In summary, we have shown that excess noise generated by Pb-salt lasers can be classified according to high frequency content and we have indicated a possible explanation for each noise category. We have shown the spatial dependence of the TDL modes and the IF noise associated with these modes which indicated that the semiconductor lasers tested sometimes emit simultaneously in two or more lasing filaments. Results presented show that multiple filaments are at times involved in the mechanism which generates excess noise although the mere presence of multiple filaments does not guarantee that excess noise will be present.

Throughout the results presented, we have shown the similarity between the $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ semiconductor lasers and the results of other workers on GaAs semiconductor lasers. Although these two materials are different and emit radiation at different wavelength regions, the similarities are such that the techniques used to control the excess noise in the GaAs semiconductor lasers may also work on the Pb-salt lasers.

REFERENCES

1. Savage, M.; Augeri, R.; and Peyton, B. J.: Application of Tunable Diode Lasers to Infrared Heterodyne Spectroscopy. Second Conference on Lasers and Applications (Orlando, Fla.), Dec. 17-21, 1979.
2. Eng, R. S.; Mantz, A. W.; and Todd, T. R.: Improved Sensitivity of Tunable-Diode Laser Open Path Trace Gas Monitoring System. Appl. Opt., 18, 3438, 1979.
3. Eng, R. S.; Mantz, A. W.; and Todd, T. R.: Low Frequency Noise Characteristics of Pb-Salt Semiconductor Lasers. Appl. Opt., 18, 1088, 1979.
4. Paoli, T. L.; and Ripper, J. E.: Coupled Longitudinal Mode Pulsing in Semiconductor Lasers. Phy. Rev. Lett., 22, 1085, 1969.
5. Paoli, T.; and Ripper, J.: Observation of Intrinsic Quantum Fluctuations in Semiconductor Lasers. IEEE J. Quant. Elect., QE-6, 325, 1970.
6. Bogatov, A. P.; Eliseev, P. G.; Logginov, A. S.; Manko, M. A.; and Senatorov, K.: Study of the Single-Mode Injection Laser. IEEE J. Quant. Elect., QE-9, 392, 1973.
7. Eliseev, P. G.; and Shuikin, N. N.: Single Mode and Single Frequency Injection Lasers. Sov. J. Quant. Electron, 3, 181, 1973.
8. Bachert, H.; Eliseev, P.; Maneo, V.; Stremov, S.; Raab; and Thay, C.: Multimode Operation and Mode-Locking Effect in Injection Lasers. IEEE J. Quant. Elect., QE-11, 507, 1975.
9. Kobayashi, T.; Takanashi, Y.; and Fusukana, Y.: Reduction of Quantum Noise in Very Narrow Planer Stripe Lasers. Jap. J. of Appl. Phys., 17, 535, 1978.
10. Harward, C. N.; and Hoell, J. M.: Optical Feedback Effects on the Performance of $Pb_{1-x}Sn_xSe$ Semiconductor Lasers. Appl. Opt., 18, 3978, 1979.
11. Broom, R. F.: Self-Modulation at Gigahertz Frequencies of a Diode Laser Coupled to an External Cavity. Electronic Letters, 5, 571, 1969.
12. Broom, R. F.; Mohn, E.; Risch, C.; and Salathe, R.: Microwave Self-Modulation of a Diode Laser Coupled to an External Cavity. IEEE J. Quant. Elect., QE-6, 328, 1970.
13. Ho, P. T.; Glasser, L. A.; Ippen, E. P.; and Haus, H. A.: Picosecond Pulse Generation With a CW GaAlAs Laser. Appl. Phys. Lett., 33, 241, 1978.
14. Risch, C.; and Voumard, C.: Self-Pulsations in the Output Intensity and Spectrum of GaAs-AlGaAs CW Diode Lasers Coupled to a Frequency Selective External Optical Cavity. J. of Appl. Phys., 48, 2083, 1977.

15. McCumber, D. E.: Intensity Fluctuations in the Output of CW Laser Oscillators I. Phys. Rev., 141, 306, 1966.
16. Haug, H.; and Haken, H.: Theory of Noise in Semiconductor Laser Emission. Z. Physik, 204, 262, 1967.
17. Armstrong, J. A.; and Smith, A. W.: Intensity Fluctuations in GaAs Laser Emission. Phy. Rev., 140, A155, 1965.
18. Smith, A. W.; and Armstrong, J. A.: Intensity Noise in Multimode GaAs Laser Emission. IBM Res. Develop., 10, 225, 1966.
19. Guekos, G.; Jackel, H.; and Schmid, K. F.: Excess Noise in the Radiation From CW D. H. GaAlAs-Diode Lasers. Elect. Lett., 12, 64, 1975.
20. Jackel, H.; and Guekos, G.: High Frequency Intensity Noise Spectra of Axial Mode Groups in the Radiation From CW GaAlAs Diode Lasers. Opt. and Quan. Elect., 9, 233, 1977.
21. Guekos, G.; and Strutt, M. J.: Laser-Light Noise and Current Noise of GaAs CW Laser Diodes. IEEE J. Quant. Elect., QE-6, 423, 1970.
22. Ito, T.; Machida, S.; Nawata, K.; and Regami, T.: Intensity Fluctuations in Each Longitudinal Mode of a Multimode AlGaAs Laser. IEEE J. Quant. Elect., QE-13, 574, 1977.
23. Kingsley, T.; and Finner, G.: Stimulated Emission From P-N Junctions. Proc. 3rd International Congr., Quantum Electron. (Paris), Dunod, 1883, 1964.
24. Hatz, J.: Some Effects on Material Inhomogeneities on the Near Field Pattern of GaAs Diode Lasers. Phy. Status Solidi, 28, 233, 1968.
25. Deutsch, C.; and Hatz, J.: Observation of Mode Coupling in GaAs Lasers. J. Appl. Math. Phys. (ZAMP), 18, 599, 1967.

TABLE I.- PEAK HEIGHTS OF MEMBERS OF LONGITUDINAL MODE
FAMILIES AT LASER CURRENTS FROM 1.337 TO 1.345 AMPS

Laser current, amps	Mode									
	1	2	3	4	5	6	7	2 + 3 + 6	1 + 4 + 5 + 7	Sum overall
1.337	15	0	0	19	12	0	12	0	58	58
1.338	16	0	0	19	12	0	12	0	59	59
1.339	17	0	0	19	11	0	12	0	59	59
1.340	16	1	1	19	8	0	13	2	56	58
1.341	7	7	16	8	3	4	12	27	30	57
1.342	6	8	18	6	3	4	12	30	27	57
1.343	6	8	19	5	2	5	13	32	26	58
1.344	5	8	20	5	2	5	13	33	25	58
1.345	5	8	21	4	2	5	13	34	24	58

TABLE II.- PEAK HEIGHTS OF MEMBERS OF LONGITUDINAL MODE
FAMILIES AT LASER CURRENTS FROM 1.7895 TO 1.8005 AMPS

Laser current, amps	Mode								
	1	2	3	4	5	6	1 + 3 + 5	2 + 4 + 6	Sum overall
1.7895	5	8	9	14	10	8	24	30	54
1.7905	5	11	8	14	11	6	24	31	55
1.7915	4	13	8	13	12	3	24	29	53
1.7925	3	15	8	13	13	2	24	30	54
1.7935	3	17	8	13	13	1	24	31	55
1.7945	10	5	10	15	5	7	25	27	52
1.7955	11	4	11	18	5	8	27	30	57
1.7965	12	3	11	19	4	7	27	29	56
1.7975	12	3	11	20	4	7	27	30	57
1.7985	13	3	11	19	4	6	28	28	56
1.7995	15	2	11	20	3	5	29	27	56
1.8005	16	3	10	21	4	4	30	28	58

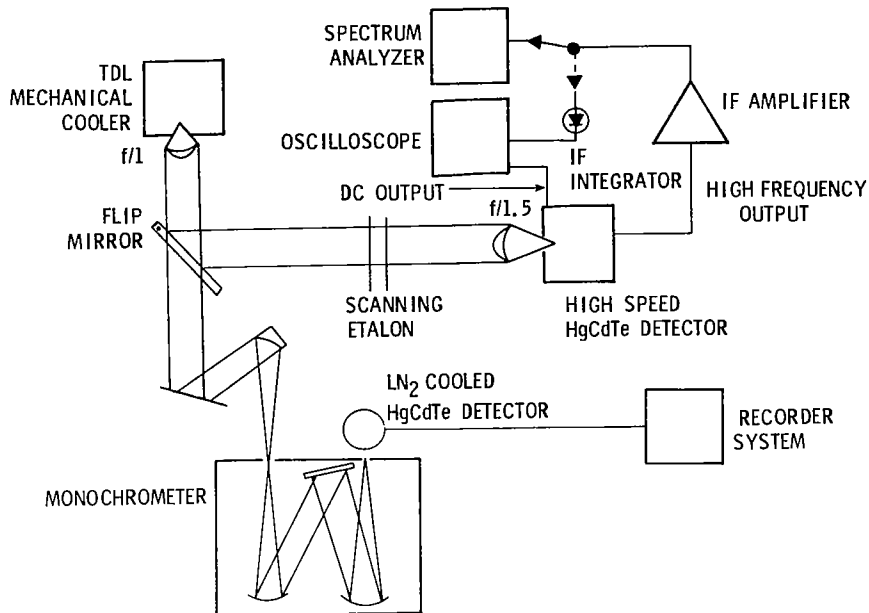


Figure 1.- TDL characterization facility.

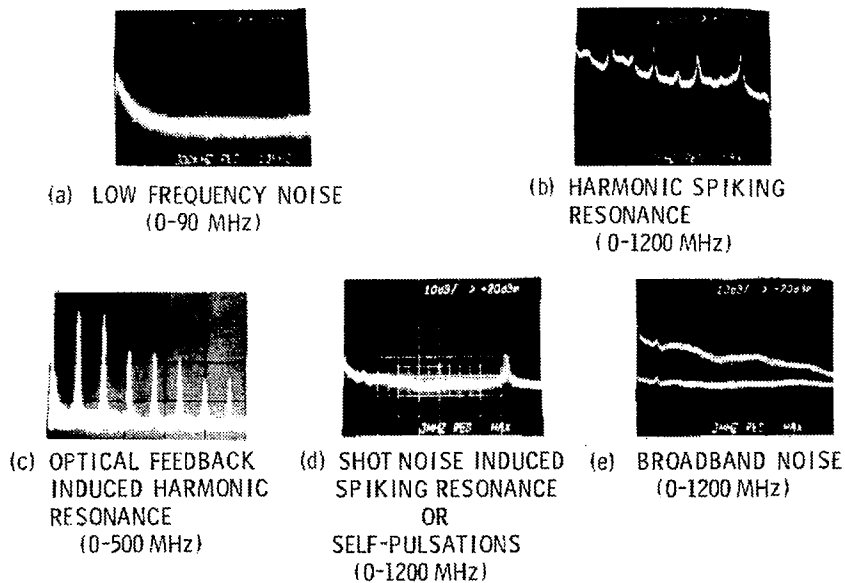


Figure 2.- Categories of excess noise.

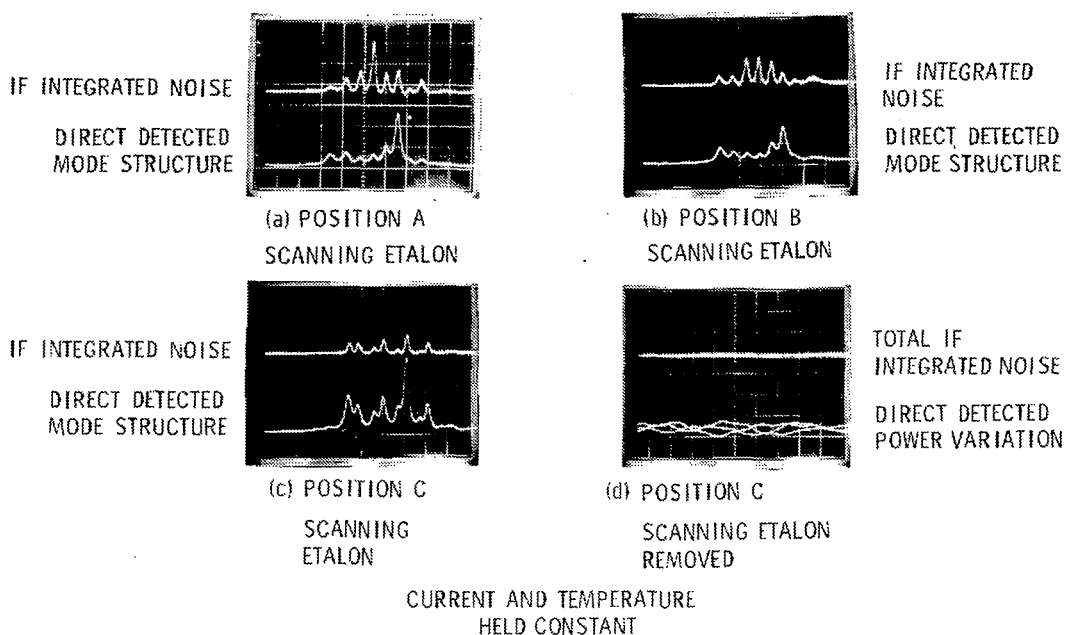


Figure 3.- Spatial dependence of TDL mode power and excess noise; noise amplitude stabilization.

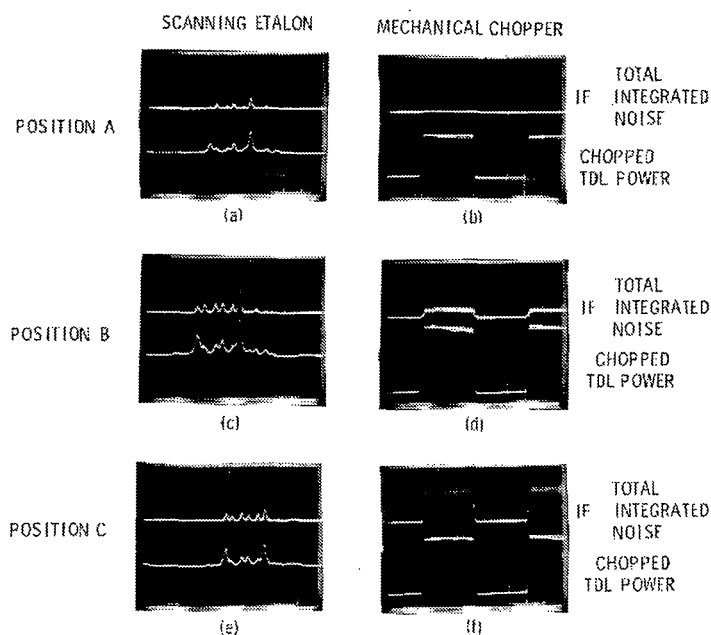


Figure 4.- Spatial dependence of TDL mode power, excess noise, and noise amplitude stabilization.

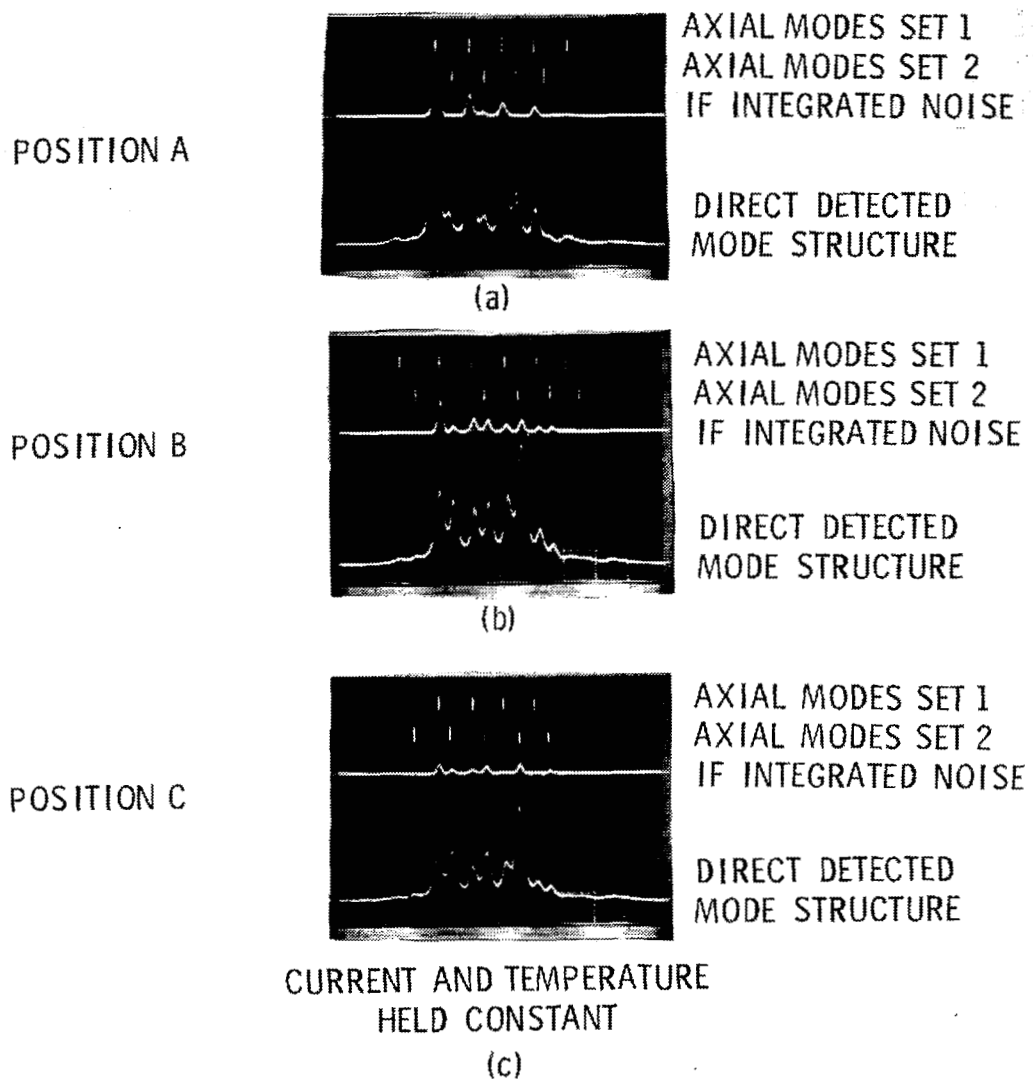


Figure 5.- Spatial dependence of TDL mode power and excess noise.

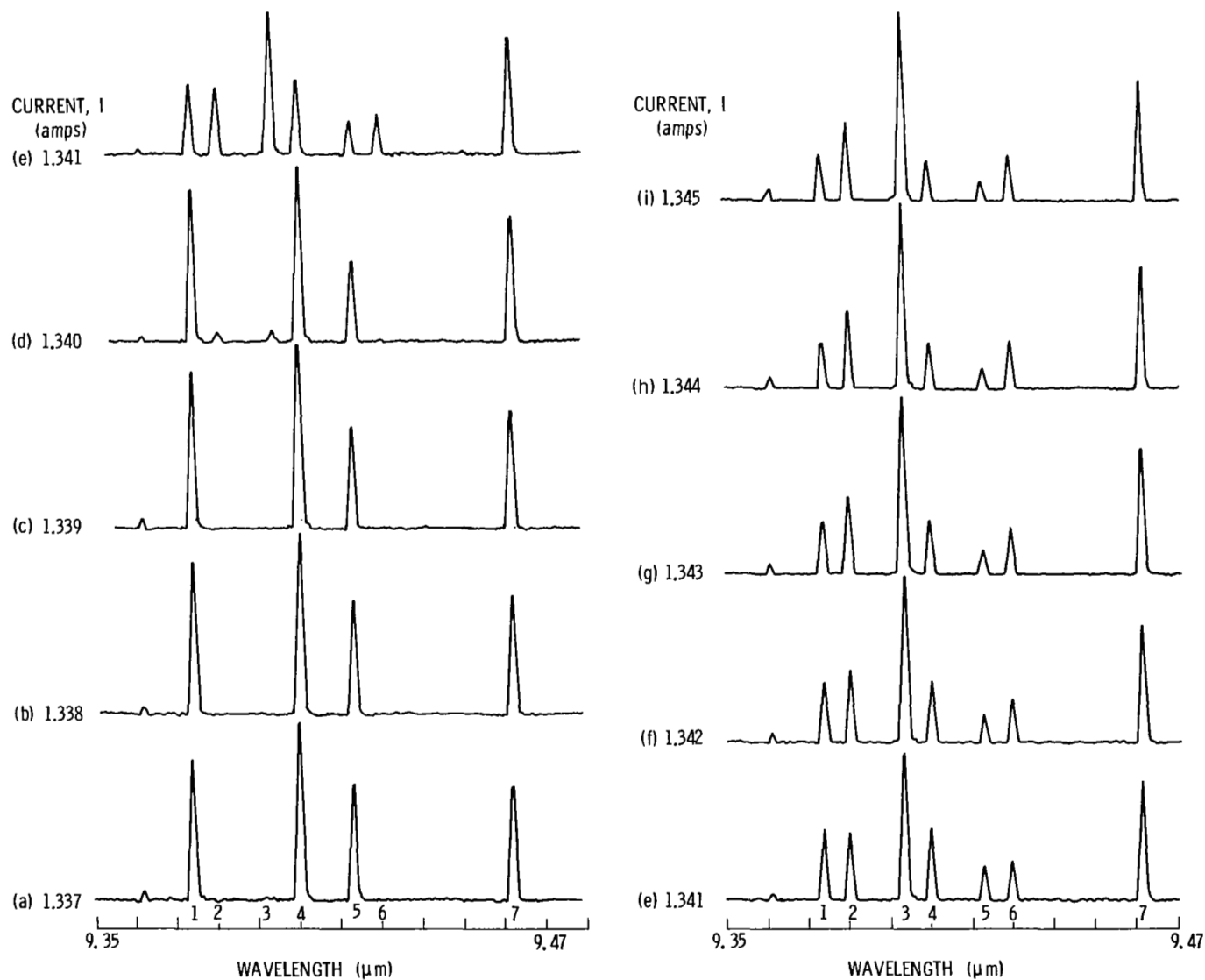


Figure 6.- Current dependence of TDL frequency modes for $I = 1.337$ to 1.345 amps.

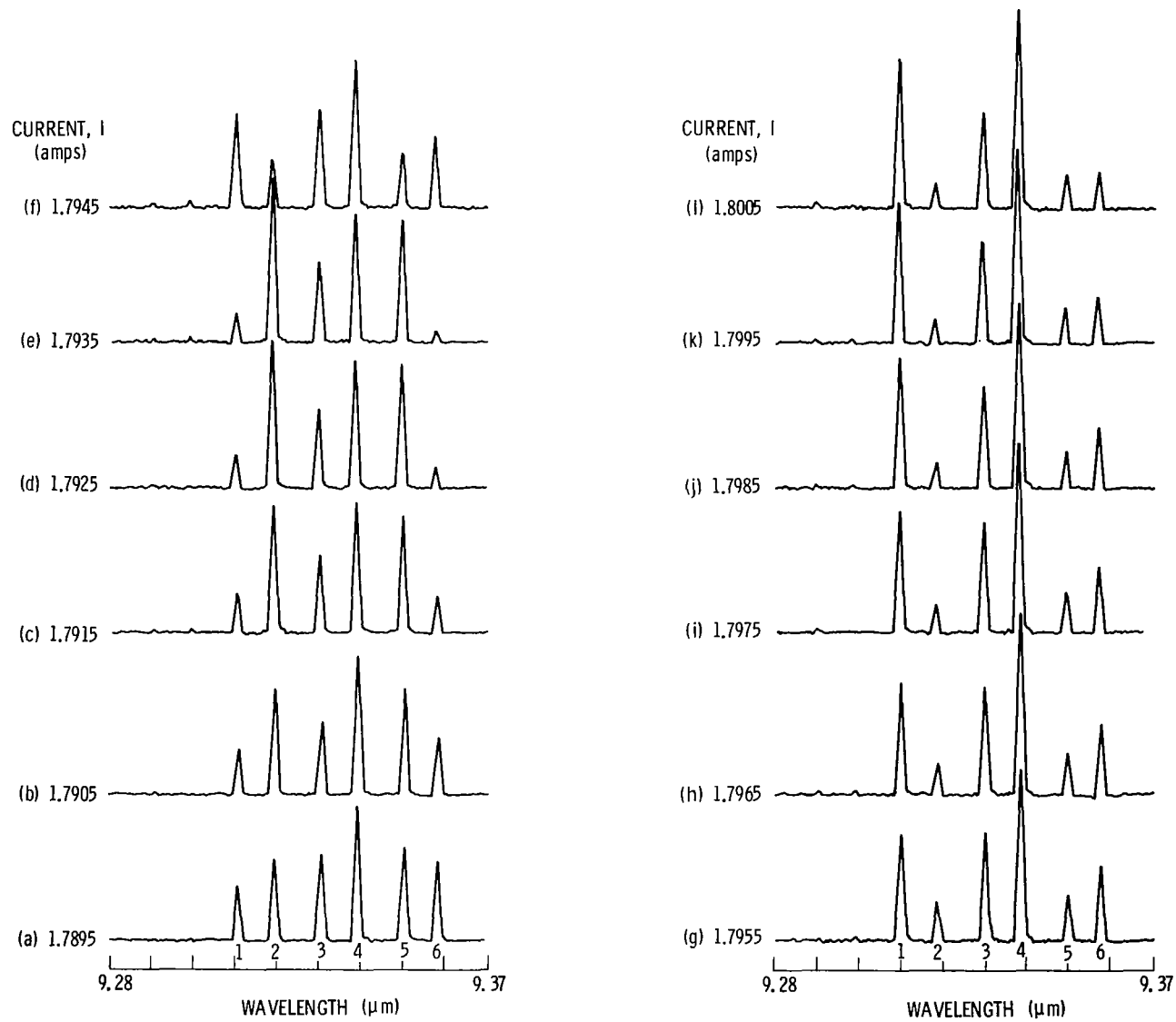
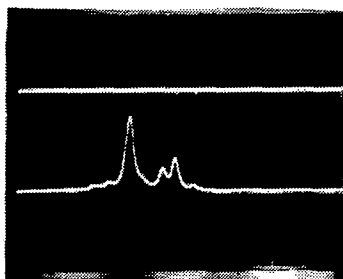


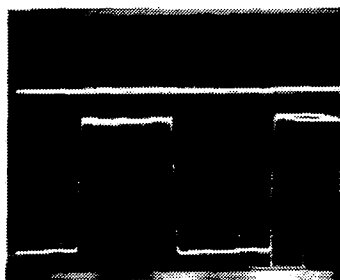
Figure 7.- Current dependence of TDL frequency modes for $I = 1.7895$ to 1.8005 amps.

SCANNING
ETALON



IF INTEGRATED NOISE
DIRECT DETECTED
MODE STRUCTURE

NO
SCANNING
ETALON



TOTAL IF INTEGRATED
NOISE
CHOPPED DIRECT
DETECTED TDL
OUTPUT

Figure 8.- TDL mode power and total power in absence of excess noise.

FINE WAVELENGTH ID FOR TUNABLE LASER

LOCAL OSCILLATORS*

M.G. Savage and R.C. Augeri

Eaton Corporation AIL Division, Melville, New York 11747

SUMMARY

This paper describes a wavelength ID device which consists of an electronic servo-controlled Fabry-Perot etalon. Measurements performed with a CO₂ laser show that the etalon has a finesse $F > 30$ which is maintainable for several days. These tests also demonstrate that the etalon system is capable of resonance frequency stability $\frac{\Delta\nu}{\nu} \approx 2 \times 10^{-7}$ during similar time periods. With currently available mirror coatings, this level of performance is achievable over an optical bandwidth $\Delta\lambda = 3 \mu\text{m}$ centered at $\lambda = 10 \mu\text{m}$.

INTRODUCTION

With the increased concern about possible damage to the atmosphere by chemical pollution, much interest has arisen in the chemistry of the stratosphere. Several important constituents which are present in only trace amounts, such as ozone O₃ (reference 1), chlorine monoxide ClO (references 2 and 3), and the various chlorofluorocarbons (reference 4) have prominent absorption-emission spectra in the 8 to 12 μm wavelength region. Remote sensing of these and many other gases can be accomplished with infrared heterodyne spectrometers using tunable diode lasers (TDL) as local oscillators (LO) (references 5 and 6).

TDL's exhibit narrow spectral emission widths which are piecewise-continuously tunable over more than 100 wavenumbers (references 7, 8, 9, and 10). While such tunability greatly increases the spectral coverage obtainable with a single-TDL spectrometer, it introduces a degree of uncertainty in the LO frequency which is not present when a low-pressure gas laser is used. In studies of spectrally narrow features such as stratospheric gas absorption lines, this LO frequency uncertainty can have serious effects on measurement accuracy (reference 11). It is often necessary to determine the LO frequency with a resolution $\frac{\Delta\nu}{\nu} = 1.8 \times 10^{-7}$ (about 5 MHz at $\lambda = 10.6 \mu\text{m}$).

*This work was supported by NASA, Langley Research Center, Hampton, Virginia 23665.

WAVELENGTH ID TECHNIQUES

Measurements of TDL output parameters such as wavelength versus drive current have generally shown large temporal variability rendering wavelength ID via drive current calibration nearly useless (references 7 and 10). A successful means of accurate wavelength determination will rely as little as possible on the stability of the TDL.

The use of a reference gas absorption cell is capable of providing the necessary wavelength resolution if a suitable absorber can be found. (It is generally impossible to use the target gas line as a reference because of the conflicting requirements of a low absorption coefficient for viewing through kilometers of atmosphere and a high absorption coefficient for viewing through centimeters of test cell.) To be used successfully, the absorption spectrum of the reference gas must be determined with the same resolution as that desired of the TDL's wavelength. Since there are many atmospheric gases of interest, high resolution spectra of many potential reference absorbers must be generated. In addition, some knowledge of the TDL's wavelength-current tuning rate is required by this technique because it is unlikely that the reference line frequency will exactly coincide with the target line frequency. The larger this separation is, the higher the accuracy with which the TDL's tuning rate must be known to maintain a given level of LO frequency accuracy. Also, the above calibration technique relies on an ageless TDL whose tuning parameters do not change.

Another technique which is capable of providing the desired wavelength ID accuracy uses a wideband detector such as a metal-oxide-metal (MOM) diode to mix the outputs of a TDL and a stabilized low-pressure gas laser. The resultant heterodyne beat note can be measured with very high precision. MOM diodes designed for operation near $\lambda = 10 \mu\text{m}$ can have IF bandwidths in excess of 30 GHz (1 cm^{-1}) allowing continuous coverage of the spectral intervals between CO_2 laser lines (reference 12). While these detectors do not have the necessary sensitivity to perform the desired atmospheric measurements, they are well suited to the wavelength ID function. This technique is limited, however, to those spectral regions where suitable gas laser oscillators are available.

Another device capable of high wavelength resolution is the Fabry-Perot etalon (FPE), which is an optically resonant cavity formed by two parallel plane partial reflectors. Mirror coatings presently available can allow an FPE to operate with a finesse $F \geq 30$ throughout a spectral interval $\Delta\lambda \approx 3 \mu\text{m}$ near $\lambda = 10 \mu\text{m}$. Until recently these devices were not stable enough to permit their potentially high resolution to be used for TDL wavelength calibration. An etalon has been designed and constructed which overcomes these stability problems, maintaining a finesse $F = 33 \pm 3$ for a period of 2 days, and showing resonance frequency stability of $\pm 10 \text{ MHz}$ over 4 hours. The design is capable of even longer term stability.

HIGH RESOLUTION TUNABLE FPE

The etalon employs a capacitance-bridge servo technique to maintain mirror parallelism and to provide continuous passband tuning (references 13, 14, and 15). The FPE uniquely uses a Cer-Vit mirror spacer as the primary length standard for the optical cavity. (Cer-Vit is manufactured by Owens-Corning Fiberglas Corporation.) The thermal stability of the spacer reduces the need for precise temperature control of the etalon to a manageable level. The mirrors which form the optical cavity are attached to piezoelectric elements (PZT) which can accurately vary the cavity length and, therefore, spectrally tune the etalon.

The mirrors, which form the etalon shown schematically in Figure 1, consist of multilayer dielectric coatings deposited on 5-cm-diameter ZnSe substrates. They have a reflectivity $R \geq 0.90$ over the spectral interval $8.9 \mu\text{m} \leq \lambda \leq 11.2 \mu\text{m}$. The substrates are antireflection coated on one side and have small wedge angles. The plates are matched in flatness to better than $\lambda/175$ at $\lambda = 10.6 \mu\text{m}$. The spacer is a 4-cm-long cylinder whose surfaces were ground flat and parallel to nearly $\lambda/1000$ at $\lambda = 10.6 \mu\text{m}$.

Gold has been sputtered onto the mirror and spacer surfaces in a pattern shown in Figure 2. When the mirrors are positioned close to the spacer, the opposing gold pads form parallel plate capacitors. The pads on the spacer are slightly oversized to facilitate mirror-spacer alignment. The gaps between the mirrors and the spacer are manually preset to a nominal 50 μm .

Each mirror of the FPE is independently controlled by an electronic servo mechanism. The two pairs of orthogonal capacitors, arbitrarily designated $X_1 - X_2$ and $Y_1 - Y_2$, are used to maintain mirror-spacer parallelism, thereby also assuring mirror-mirror parallelism. This is accomplished by balancing the pairs in separate capacitance bridges. As the mirrors drift out of parallel, the resulting bridge voltages are sensed and used to drive the PZT's supporting the mirrors. Cavity length control and therefore resonance frequency control is performed by balancing the two unpaired Z-capacitors against fixed-value reference capacitors in two other bridges. The frequency stability of the etalon is dependent on the stability of the fixed reference capacitors and the Cer-Vit spacer.

Since the FPE mirrors are individually controlled, six high-resolution capacitance bridges are required to form the servo system. A block diagram of the electronics used to control one of the mirrors is shown in Figure 3 (reference 13). Each of the three bridges has a balance control to compensate for stray capacitances arising from either the etalon structure or the servo itself. Spectral tuning of the FPE is accomplished by combining the output of the Z-bridge with an adjustable offset voltage; the combination then is used to drive the PZT's. This tuning method is readily adaptable to digital control and allows continuous spectral coverage within limits set by the maximum PZT motion.

Changes in the pad gap ℓ are related to changes in the capacitance C according to the expression:

$$\frac{d\ell}{\ell} = \frac{dC}{C} \quad (1)$$

(Equation (1) assumes that there is a negligible area difference between the pairs of capacitors.) Also, changes in the cavity length L are related to changes in the pad gaps by:

$$d\ell = 1/2 dL \quad (2)$$

The factor of 1/2 arises when both plates are translated by the same distance dL . For one plate fixed, the relation reduces to $d\ell = dL$. To achieve the desired wavelength resolution of 1.8×10^{-7} , it can be shown from the etalon resonance condition that:

$$\frac{dC}{C} = 9 \times 10^{-8} \frac{L}{\ell} \quad (3)$$

For the case in which $L = 4$ cm and $\ell = 50$ μm , the necessary capacitance bridge resolution corresponding to an FPE resonance frequency resolution of 5 MHz at $\lambda = 10.6$ μm is $\frac{dC}{C} = 7.2 \times 10^{-5}$.

Stabilization of an FPE with such high precision requires special consideration of temperature and pressure effects. Since the spacer is the length standard of the cavity, it must be dimensionally stable to the same degree as the desired frequency resolution. Since Cer-Vit has an expansion coefficient of

$0.12 \times \frac{10^{-6} \text{ cm}}{\text{cm } ^\circ\text{C}}$, temperature stability of about 1.5°C is necessary to achieve the

desired result. This etalon is mounted in a vacuum chamber to eliminate the effects of atmospheric pressure variations on the dielectric constant of the capacitor gaps and the index of refraction of the air between the mirrors.

Measured Performance

The measurements of the FPE finesse and frequency stability at $\lambda = 10.6$ μm were performed using the setup shown in Figure 4. The collimated output of the HeNe laser was directed through the etalon to a pyroelectric detector. As the spacing between the mirrors is varied by the PZT's, peaks in the output of the detector occur which correspond to cavity length changes of $\lambda/2$ or 0.3164 μm . The voltage expansion coefficient of the PZT's is then easily calculated.

Because the mirrors are not optimized for operation at visible wavelengths, the peaks that are observed are not very sharp, introducing an error of about 10 percent into the measurement of the expansion coefficient. The output of a waveguide CO₂ laser operating at $\lambda = 10.588 \mu\text{m}$ was then directed through the etalon with an FPE passband tuned across the laser output by PZT voltage adjustment. The measured transmission versus length change can be used to estimate the finesse because the cavity length is known. Figure 5 shows two curves of calculated passband shape for two values of finesse and PZT expansion coefficient. Also shown are data points measured on two occasions two days apart.

The measurement of the frequency stability of the FPE was performed using the same optical setup by tuning the FPE so that a half-transmission point of a passband overlapped the CO₂ laser output. The detected laser power transmitted by the etalon was monitored by a strip-chart recorder. Adjusting the FPE to a passband skirt increases its frequency discrimination. The long-term frequency stability of the laser is believed to be better than 3 MHz and its output power was continuously monitored. The FPE was mounted in a vacuum chamber, but its temperature was not actively controlled.

Figure 6 shows the measured resonance frequency stability of the etalon. The dashed curve shows that the frequency deviation was less than 10 MHz over a period of 4 hours in the early afternoon. The solid curve shows a continuous drift through the evening. This drift is caused by a decrease in ceramic reference capacitor temperature as the ambient lab air cooled. Measurements of spacer temperature during this period indicated a stability of better than 1°C. The figure clearly shows that the FPE is spectrally stable to an unprecedented degree provided sufficient reference capacitor thermal stability is achieved. An improved air gap reference is being developed.

WAVELENGTH ID WITH AN FPE

The schematic diagram of a possible TDL wavelength calibration system using a high-resolution FPE is shown in Figure 7. The output of a TDL is first prefiltered to isolate a single frequency mode. Part of this power is directed through the FPE to a detector whose output is then fed back to control the TDL temperature or drive current. Depending on the available TDL power, FPE parameters, and calibration detector sensitivity, it may be necessary to impress a slight frequency modulation on the TDL output to enhance calibration accuracy. This modulation would have no significant effect on the laser's performance as an LO.

It is absolutely necessary to prefilter the TDL output; this can be done with a small monochromator. The removal of extraneous TDL frequency modes is required by the spectrometer to avoid modal self-beating effects and possible IF ambiguity (reference 10). The bandpass of this monochromator can be as large as a TDL mode width ($\sim 1 \text{ cm}^{-1}$), if necessary, but some additional filtering may be required in the calibration system to position the TDL output within a particular high-resolution FPE passband.

High-finesse, long-cavity etalons need special attention for their beam quality requirements. Both beam divergence and variations of angle of incidence must be minimized to extract the best FPE system performance. For the etalon previously described, the beam divergence half-angle should be less than a few tenths of a degree to avoid excessive resonance frequency shifts and finesse degradation. Since the etalon is used at normal incidence, its sensitivity to changes in incidence angle is somewhat reduced, but these changes should still be kept below several tenths of a degree to avoid too much frequency shift.

CONCLUSIONS

For tunable diode lasers to realize their potential in high-resolution heterodyne spectroscopy, some means of accurately determining their emission frequency is required. In general, characteristics of a TDL are not sufficiently stable over the long term to allow their one-time calibration. Several potential calibration techniques have been discussed which have varying degrees of flexibility and accuracy. Any of these methods could be incorporated into a spectrometer as a subsystem. A Fabry-Perot etalon has been developed which is capable of calibrating a TDL output with a resolution of about 5 MHz at wavelengths near 10 μm and measured data has been presented. This device can form the basis of a highly flexible and accurate wavelength ID system.

REFERENCES

1. R. A. McClatchey, W. S. Benedict, S. A. Clough, D. E. Burch, R. F. Calfee, K. Fox, L. S. Rothman, and J. S. Garing, AFCRL Atmospheric Absorption Line Parameters Compilation, AFCRL-TR-73-0096 (1973).
2. R.T. Menzies, J. S. Margolis, E. D. Hinkley, and R. A. Toth, Appl Opt 16, 523 (1977).
3. R. S. Rogowski, C. H. Bair, W. R. Wade, J. M. Hoell, and G. E. Copeland, Appl Opt 17, 1301 (1978).
4. J. M. Hoell, C. H. Bair, C. Harward, and B. Williams, Geophys Res Lett 12, Dec 1979.
5. M. A. Frerking and D. J. Muehlner, Appl Opt 16, 526 (1977).
6. R. T. Ku and D. L. Spears, IEEE J-QE QE-13, 72D (1977).
7. G. A. Antcliffe and S. G. Parker, J. Appl Phys 44, 4145 (1973).
8. K. J. Linden, K. W. Nill, and J. F. Butler, IEEE J-QE QE-13, 720 (1977).
9. W. Lo, G. P. Montgomery, Jr., and D. E. Swets, J. Appl Phys 47, 267 (1976).
10. M. G. Savage, R. C. Augeri, and B. J. Peyton, NASA Contractor Report 159121.
11. R. K. Seals and B. J. Peyton, Proceedings of the International Conference on Environmental Sensing and Assessment, Vol 1, p 10-4, Sept 1975.
12. A. Sanchez, C. F. Davis, Jr., C. C. Liu, and A. Javan, J. Appl Phys 49, 5270 (1978).
13. T. R. Hicks, N. K. Reay, and R. J. Scadden, J. Phys E: SCI INSTR 7, 27, (1974).
14. T. R. Hicks, N. K. Reay, and C. L. Stephens, Astron and Astrophys 51, 367 (1976).
15. R. V. Jones and J. C. S. Richards, J. Phys E: SCI INSTR 6, 589 (1973).

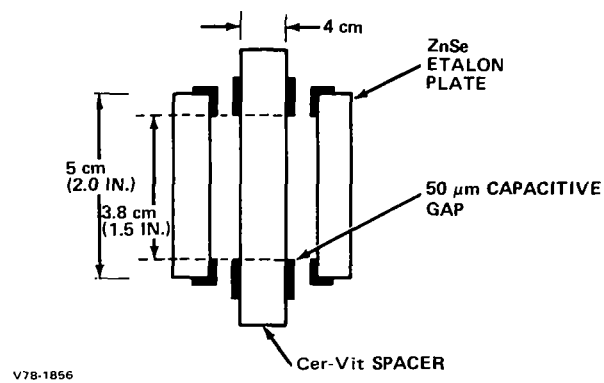


Figure 1.- Fabry-Perot etalon schematic.

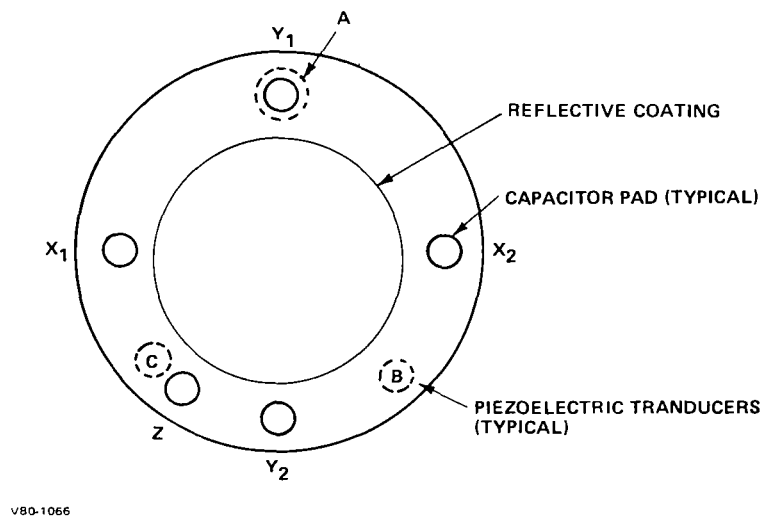


Figure 2.- Plan view of etalon plate showing capacitance pads and piezoelectric transducers.

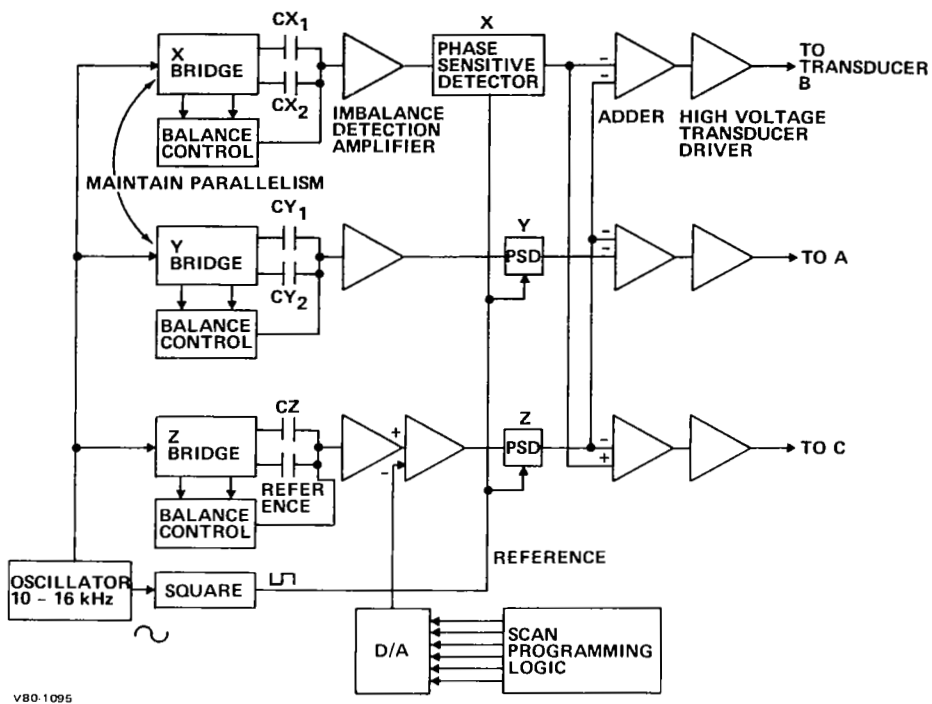


Figure 3.- Capacitance bridge electronics block diagram.

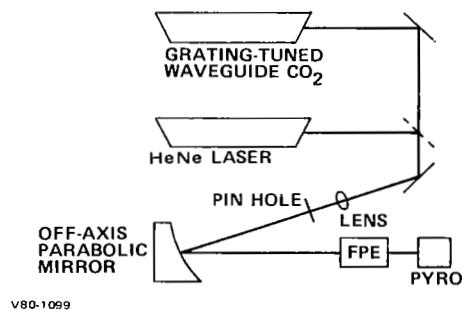
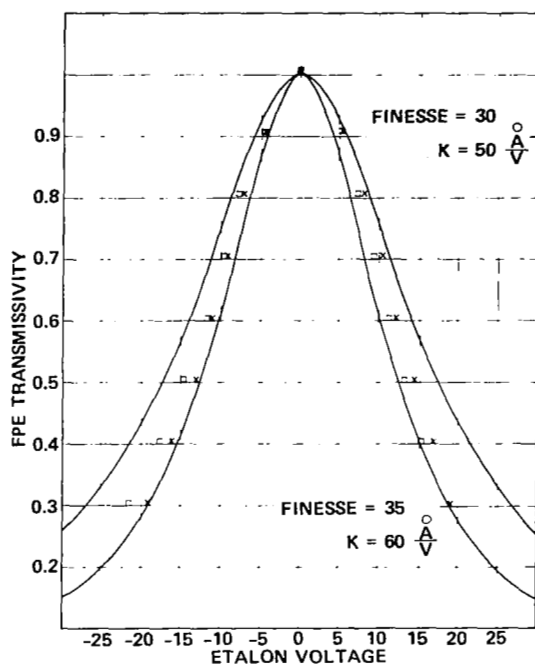


Figure 4.- Optical setup for measurement of FPE finesse and stability.



V80-1096

Figure 5.- Measured high-resolution FPE finesse.

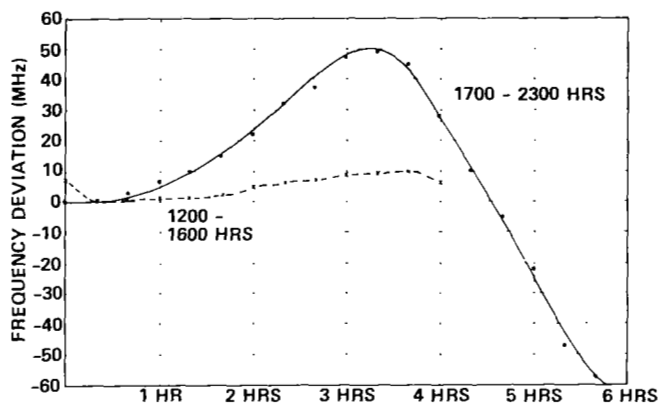
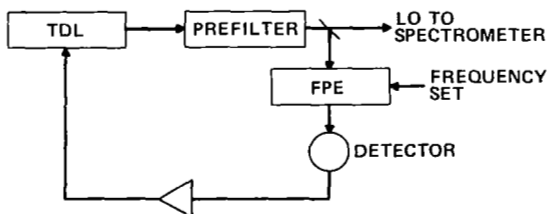


Figure 6.- Measured FPE resonance frequency stability.



V80-1098

Figure 7.- TDL wavelength ID system.

A REVIEW OF THE NASA/OAST CRYOGENIC COOLERS TECHNOLOGY PROGRAM

J. G. Lundholm, Jr.
NASA Headquarters

Allan Sherman
Goddard Space Flight Center

SUMMARY

An ongoing NASA research and technology program provides for the development of low and ultra-low temperature cryogenic cooler systems for future space missions. These include mechanical, solid cryogen, gas adsorption, superfluid helium, helium-3, and magnetic (adiabatic demagnetization) coolers.

Operating lifetimes required vary from a few weeks for a Shuttle/Spacelab mission to as long as nine years for missions to the outer planets. Temperature requirements vary from tens to tenths of kelvin. At the higher temperature, cooling loads for detectors, instruments and associated shields may be as high as 15 watts. At the ultra-low temperatures, cooling loads must be limited to tens of microwatts. All systems must survive launch loads and then function in the zero gravity environment of space.

The requirements of long-lifetime, zero-g operation and a variety of temperatures and cooling loads present complex technological problems. Much progress is being made in solving these problems.

A large number of projected NASA missions will utilize instruments that require the cryogenic coolers. These include earth resources, high energy astronomy, infrared astronomy, and gravitational physics.

INTRODUCTION

The utilization and application of low and ultra-low temperature devices and systems have been increasing at a rapid pace. This is true both for NASA and DOD space flight programs as well as many research laboratories outside of NASA. Planning for future NASA missions in the next decade indicates increasing use of low and ultra-low temperature sensors and instruments.

Although some types of coolers have been flown on a few missions to date, none of the coolers can operate for the desired lifetime in zero-g and at the extremely low temperatures required in some cases. For this reason, NASA is conducting a broad-range cooler Research and Technology (R&T) program. This will assure that the required technology will be available in time to support future missions.

The objectives of this paper are to: (1) acquaint the reader with the present cooler program managed by the Office of Aeronautics and Space Technology Information Systems Office, and (2) describe changes in program thrust from that discussed in previous papers (Ref. 1 and 2). The next section will discuss the general requirements for low and ultra-low temperature and techniques that can be employed to achieve these temperatures at expected heat loads. This part of the paper will be followed with a discussion of the NASA Centers involved in the R&T efforts and the type of cooler systems on which they are focusing their effort. A brief description and thrust of R&T effort will then be given of each cooler type.

COOLING REQUIREMENTS

The expected heat loads that the coolers must be capable of handling are listed in Table I. It should be noted that the heat loads must be kept extremely small (microwatts) for operation at the ultra-low temperatures. A comparison of the data in Tables I and II allows one to determine the type of cooler most likely to be used for a particular mission category. Earth observation missions with instruments using arrays of IR sensors looking at the earth will require mechanical coolers which can supply a few watts cooling capacity and temperatures of a few tens of kelvin. Infrared astronomy missions such as IRAS (Infrared Astronomy Satellite) will use superfluid helium which can provide temperatures down to 1.5K with tens to a few hundred milliwatts cooling loads. Infrared bolometers may be required to be cooled to 0.3K using a helium-3 cooler to obtain data at longer wavelengths. For instance, the sensitivity of a bolometer continues to increase as its operating temperature is lowered. In the longer term, relativity or gravitational physics missions, which will attempt to determine the existence of gravitational waves, may require temperatures of a few millikelvin. These low temperatures can only be achieved with an adiabatic demagnetization refrigerator. In addition to mission requirements, special coolers will be required to cool experimental facilities to perform various basic science experiments in space. One of the presently approved Spacelab experiments will measure fundamental properties of superfluid helium (He II) under zero-g conditions. These are measurements on a "quantum fluid" which can only be performed under zero-g conditions. Tables I and II clearly show that different types of coolers must be available in order to meet the varying requirements (temperature, heat load/operating lifetime) placed on the cooler by different missions.

PROGRAM MANAGEMENT STRUCTURE

Because of the diversity of the cryogenic cooler program, three NASA Centers are involved in performing the Research and Technology efforts. The major areas of effort by Centers are as follows:

Goddard Space Flight Center (GSFC)

- o Long-Lifetime Mechanical Coolers
- o Solid Cryogen Coolers

Ames Research Center (ARC)

- o Helium-3 Coolers
- o Adiabatic Demagnetization Coolers
- o Long-Life Superfluid Helium Cryostats

Jet Propulsion Laboratory (JPL)

- o Radiant Coolers
- o Adsorption Coolers
- o Magnetic Coolers
- o Spacelab Experiment - Measure Basic Properties of Superfluid Helium in Zero-g

The cooler program is managed by the NASA Headquarters Office of Aeronautics and Space Technology (OAST). Coordination is maintained with the NASA Headquarters program offices, which will be the eventual users of the new technology. This includes the Office of Space Science (OSS) and the Office of Space and Terrestrial Applications (OSTA). The cooler program activity is also reviewed in coordination meetings with DOD.

CRYOGENIC COOLER TYPES

Radiant Coolers

Radiant coolers achieve cryogenic temperatures by radiating the instrument loads to cold space. Typical operating temperatures for present radiant coolers are about 100K with a capacity on the order of 30 mW and the sensor heat dissipation but a fraction of the total. A detailed discussion of the design of radiant coolers is given in reference 3. A recent addition to the cooler R&T program is an effort by JPL to improve the performance of radiant coolers.

The JPL concept consists of a radiator plate surrounded by fixed radiation shields. Conduction heat loads to the radiator plate are reduced by the use of fiberglass tape supports with a small cross-sectional area that are capable of supporting the plate in a lg environment, but not during launch. A caging and release mechanism is then employed for the plate. The concept could result in much larger radiant coolers in future missions with greater heat rejection capacity.

The radiant or passive cooler effort at JPL is part of an overall JPL plan to employ a combination of magnetic and gas adsorption cooler stages to pump heat from the desired low or ultra-low temperature to a sufficiently high temperature where it could be radiated to space with a radiant cooler. With no expendables, the system operating life is determined only by the reliability of the magnetic and gas adsorption cooler stages. With the JPL interest in 6 to 9 missions to the outer planets, this concept has many desirable features.

Solid Cryogen Coolers

Solid cryogen coolers utilize the heat of sublimation of cryogenics to provide the cooling. To date, temperatures down to 15K, using solid neon, are considered practical. A detailed discussion of the operation of solid cryogen coolers is in reference 4.

The specific objectives of the solid cryogen cooler Research and Technology (R&T) programs being conducted by Goddard Space Flight Center (GSFC) are capacity enhancement to 3 watt-years and extension of the usable temperature range down to 8 K using solid hydrogen. To date, the emphasis has been placed upon the latter objective.

For many missions, solid cryogen coolers have the advantages of relatively simple design and elimination of zero-g fluid management considerations. Extending the cooling capacity to 3 watt-years or more will allow for a wide range of uses, including operational missions. Extending the temperature range to 8 K may allow for the cooling of infrared detectors for earth resources satellite missions by a simpler and more compact cooler as compared to a helium cooler. This could result in lower system costs.

Temperatures of 8 K to 13 K should be able to be achieved with solid hydrogen. The subliming hydrogen is exposed to the vacuum of space in order to achieve this low temperature. Solid cryogenics used up to now include solid neon, which can provide temperatures of 13-24 K, solid nitrogen (43-63 K), solid argon (48-83 K), and solid methane (60-90 K).

A contractual effort is underway with Lockheed Research Laboratories to perform fill, boil-off and vent tests on a flight-type solid hydrogen cooler. Figure 1 shows a schematic of the test setup for this effort. Testing of the hydrogen cooler was successfully completed and the results will soon be documented. A smaller contractual effort with Beechcraft Corporation, in which a laboratory type cooler was filled with liquid hydrogen and then the hydrogen was frozen, is also now complete.

In order to achieve the 3 watt-year cooling capacity, heat leaks along the tank supports must be kept low by employing low conduction materials and a long conduction path. To reduce the radiant input, many layers of super insulation are utilized. The cool gas from subliming solid cryogen cools multiple metal barriers that are placed between the many layers of super insulation. It may be necessary to place a low temperature cryogen stage inside a higher-temperature cryogen system of solid ammonia or carbon dioxide, which is used to cool the outer thermal shield of the low temperature system. This is known as a two stage or guarded system.

As an example of a flight solid cryogen system, Figure 2 is a photograph of a two-stage cooler for the Limb Radiance Inversion Radiometer that was orbited aboard Nimbus-F. Figure 3 is a cut-away view of the two-stage cooler. The cooler, which was developed by Lockheed Research Laboratories, uses methane as the primary cryogen at 63 K and ammonia at 152 K as the secondary. With

6.4 kilograms of solid methane and 5.4 kilograms of solid ammonia, an operating lifetime of 7 months was achieved (ref. 4). This cooler also flew aboard the Nimbus-G spacecraft with similar performance.

Mechanical Coolers

Many of the future NASA missions require long-term cooling in the 5-10 watts load. A 3-5 year lifetime mechanical cooler offers a compact, low-weight system for accomplishing this goal. It is virtually the only practical system for covering this requirement. Prime mission types include long-duration gamma-ray astronomy, atmospheric research, earth resources and weather satellites. Many types of Shuttle Sortie missions (1 to 4 weeks duration) would also benefit from a reliable mechanical cooler.

Goddard Space Flight Center is the lead Center in the efforts to develop reliable mechanical coolers. The specific objective is to develop the technology for a 3-5 year lifetime cooler that will provide 5 watts of refrigeration at 65 K. After achieving this objective, efforts will be concentrated on extending this technology to the 12 K region.

The basic problem is to develop a closed cycle machine with moving parts that will run reliably for billions of cycles while unattended. Past attempts to solve this problem have used liquid or semi-liquid lubricants, dry lubricants, and "hard-on-hard" bearings with little or no lubricant. Various seal designs were also tried. To date, these approaches have not yielded a long-lifetime spaceborne cooler system.

The new approach initiated by NASA Goddard Space Flight Center hopefully eliminates the problem-causing features of the past. The key elements of the approach are: (1) the reciprocating components are driven directly with linear motors, and (2) while operating, there is no contact between the moving components and the machine housing or motor. The noncontact operation can be achieved by either magnetic or gas bearings and clearance seals. Consequently, two major contractual efforts have been funded - one to develop the magnetic bearing cooler and the other to develop the gas bearing machine.

To date, the Phase 1 design and component test activities for both the gas and magnetic bearing coolers are complete. Fabrication of the magnetic bearing design has begun, with a scheduled delivery of an engineering model cooler in January 1981. From this point on, most of the effort will be concentrated on the magnetic bearing cooler, although some component level development of gas bearings will continue.

A schematic of the Philips Laboratory magnetic bearing cooler is shown in Figure 4. The piston and the displacer are sinusoidally oscillated by the linear motors shown. These reciprocating elements are accurately positioned in their respective cylinders using active magnetic bearings. These bearings employ eddy current sensors to monitor the gap at two positions, 90° apart, around the cylinder circumference. The signals are used to adjust the currents in the four circumferential electro-magnets. The centering accuracy of these

bearings is better than 0.025 mm (0.1 mil) and they permit the efficient use of piston and displacer clearance seals which are an order of magnitude larger. Magnetic bearings support the piston on each side of the piston motor. In this manner, they equally share any radial instability force developed by the motor. A moving magnet motor design, as opposed to a moving coil, provides high reliability without flexing leads. The motor coils are hermetically sealed in metal cans to eliminate outgassing products which can affect long-term cooler stability.

The room temperature portion of the displacer is guided by two magnetic bearings. These bearings are located at the 300 K ambient temperature heat exchanger as shown. In the engineering model cooler, the displacer is driven in both directions and a restoring spring is not employed. A penalty in drive power will be paid for the simplicity of this design. A restoring "magnetic spring" may be included in future models.

The piston and displacer amplitudes and relative phase are under closed loop control. The peak-to-peak amplitude of the piston is monitored using two LVTD's. Phase control is required to maximize cooling capacity and thermal stability. The power requirement for the 5 watt, 65 K Philips cooler is about 170 W and the machine is about 70 cm long.

The quest for a reliable spaceborne mechanical cooler has been an extremely difficult one. It is hoped that the innovative approach of the Goddard Space Flight Center will result in a major breakthrough.

Adsorption Cooler

A concept for an adsorption refrigerator that is being developed by the Jet Propulsion Laboratory (JPL) is shown in Figure 5. Heat is used to compress the working gas by driving it from a zeolite adsorber. The gas is then driven through an expansion engine or a Joule Thomson device to provide the cooling. The downstream section of the cooler is maintained at near vacuum conditions by maintaining the adsorber at a low temperature through the passive radiator. When all of the working fluid ends up in the downstream adsorber, the cycle must be reversed. This is accomplished by appropriately reversing the positions (i.e., open to closed and vice versa) of the downstream and upstream heat switches and valves.

The adsorption refrigerator has the advantage of relatively few moving parts. Questionable areas are its efficiency and need for a radiant cooler. Work is progressing on the development and engineering model cooler at JPL.

Magnetic Coolers

The concept of a continuously operating magnetic cooler is also being pursued by JPL. W. A. Steyart of Los Alamos Scientific Laboratory (LASL) has published results of his work (ref. 5) on the magnetic cooler. LASL is presently cooperating with JPL to further define the concept.

The magnetic cooler can operate with efficiencies of 80% of Carnot, or greater, in the temperature range of 2 to 20 K. No high pressure seals are required. Low speed operation with a minimum number of moving parts appears possible. However, a large magnetic field of 5 to 7 Tesla is required for efficient operation.

Figure 6 shows one proposed system concept. The magnetic cooler pumps heat from the sensor load at 2.2 K to a temperature of 17 K. Additional magnetic or adsorption cooler stages would then pump the heat to a temperature of 180 K where it would be radiated to space by a large, but lightweight, radiator.

The operation of the rotating magnetic cooler is as follows: The rim of the wheel is made of porous $\text{Gd}_2(\text{SO}_4)_3 \cdot 8\text{H}_2\text{O}$, $\text{Gd}_2(\text{SO}_4)_3$, or GdPO_4 . The application of the magnetic field to the wheel in the small region shown will produce adiabatic heating. The heat is removed by the high temperature loop. As the Gd moves out of the magnetic field it drops in temperature (as much as 14 K) due to adiabatic cooling.

The rotating magnetic cooler considered by Steyart (ref. 5) allowed cycle times of much less than 1 second, that is, a wheel rotation rate of 1 to 5 revolutions per second. Circulating fluids forced through the porous wheel exchange the heat.

Stored Liquid Helium Coolers

Liquid helium (He^4) at atmospheric pressure is at a temperature of 4.2K. By pumping on the liquid helium the temperature can be lowered to 1.2 to 1.5K. At these temperatures, all other normal substances, except helium-3, are solids. Because of the low latent heat of liquid helium, high quality dewars of cryostats must be used to prevent rapid loss of the liquid.

The common isotope of helium (He^4) above 2.17K (the lambda point) is a normal liquid in all respects. Special effort must be taken to maintain thermal contact between the cryogen and the heat load. Approaches include surface tension, position retention, thermally conductive nets, and continuous acceleration by rotation of the dewar. At temperatures below 2.17K, normal helium (He I) becomes a superfluid (He II) with a thermal conductivity almost 1000 times that of copper at room temperature and almost zero viscosity. It forms a thin film on exposed surfaces. Furthermore, it moves to the warmer region which is what one would want a liquid to do if one desired an extremely uniform temperature throughout a large volume.

At present, it appears that superfluid helium has many desirable properties which make it the prime choice for cryogenic cooling in the 1.5 to 10K temperature range in systems with heat loads in the range of ten to a few hundred milliwatts. Reasonably sized cryostats can provide continuous cooling for infrared astronomy missions for periods of one year in space with expected heat loads from detectors and preamplifiers (ref. 6). With improved techniques for

installing multilayers of insulation and new designs for the inner tank supports, operating lifetimes with expected heat loads of up to three years appear possible. Consideration is being given to new designs for fiberglass tank supports. Retractable supports are also being considered. However, practical methods are needed to "fill" the hole in the super insulation blanket when the tank supports are retracted. Guarded systems using low-temperature solid-cryogen (e.g., hydrogen) or a mechanical cooler to cool the outer metal shield are another approach to achieving a 3 to 5 year life cooler.

Helium-3 Coolers

The availability of quantities of the pure isotope helium-3 (He^3) in the late 1950 time period led to its widespread use as a means of achieving lower temperatures. This is because the vapor pressure of this isotope is much higher at a given temperature than that of helium-4 (He^4). For example, at 1K, liquid He^3 (normal boiling point 3.2K) has a saturated vapor pressure 80 times larger than He^4 . By appropriately pumping on He^3 , temperatures of about 0.3K can be attained.

Ames Research Center is pursuing R&T efforts on He^3 coolers that can operate in zero-g. Because of the high cost of He^3 , it must be used in a closed-cycle mode. Special techniques must be employed to control the location of the liquid and to maintain thermal contact with the sensor or parts to be cooled.

Figure 7 is a photograph of an ARC prototype He^3 cooler. The He^3 is absorbed in a zeolite bed as it evaporates from the copper foam. To recycle, the zeolite is heated to pressurize the He^3 vapor and the copper foam is cooled to liquid helium (He^4) temperatures (probably superfluid helium at less than 2K temperature). The He^3 is then condensed onto/into the copper foam. If continuous or semi-continuous cooling is desired, multiple units might be employed with one unit being recycled while the other unit is providing temperatures as low as 0.3K.

Initial tests have shown that the unit will operate in either an upright or inverted (negative g) position. Other tests will be performed on the NASA Lear Jet and KC-135 airplanes which can provide about 30 seconds of nearly zero-g environment.

The lower temperature provided by the He^3 cooler (0.3K) as compared to a superfluid He^4 cooler (1.5K) can provide greatly increased performance for bolometer infrared detectors. In fact, the sensitivity of a bolometer IR detector will increase by a factor of 50 if cooled to 0.3K instead of 1.5K.

Joule-Thomson Cooler

Another type of cooler in which Ames Research Center (ARC) and the Jet Propulsion Laboratory (JPL) have performed preliminary tests is the Joule-Thomson expander (JTX). For this cooler, the helium may be stored under super-

critical conditions. This eliminates consideration of liquid/gas location and maintaining thermal contact in zero-g. Temperatures down to 1.4K have been achieved.

The concept has the advantage of a simplified helium storage system at pressures between 0.35 and 2.07 MPa (50 and 300 psi). Tests run on the JTX have shown sufficiently low noise levels generated by the expansion process that it will not affect the performance of infrared detectors.

Tests to date have shown that supercritical helium storage with a JTX is a viable alternative to the superfluid helium cooler for obtaining temperatures at least down to 2K with heat leaks in the tens to few hundred milliwatts range. It is being considered as one of the candidate coolers for the Shuttle Infrared Telescope Facility (SIRTF). The concept also has the advantage that there is no liquid sloshing to affect telescope pointing stability.

Adiabatic Demagnetization Cooler

The lowest practical temperature that can be produced by pumping on He³ liquid is around 0.3K, and, on He⁴ liquid is in the range of 1.2K to 1.5K. A technique by which temperatures approaching a few microkelvin has been achieved in the laboratory depends on the magneto-thermodynamic properties of complex salts. The principle of operation is as follows: the paramagnetic salt is suspended in liquid helium. A magnetic field is applied. The salt will derive energy from the magnetic field which appears as heat and raises the temperature of the salt. However, the salt is soon cooled down to the liquid helium temperature. The salt is then removed from the liquid helium and then the magnetic field is removed. The salt now cools practically instantaneously to a very low temperature.

The salt will warm up as it provides the desired cooling function. The rate at which it warms up depends on the amount of salt, the cooling load, and the degree of isolation provided by the cryostat design. After a period of time, the magnetization, cooling, and demagnetization cycle must be repeated. Multiple units might be employed whereby one is being recycled while other units are acting as coolers.

NASA Ames Research Center is performing preliminary studies on this cooling concept for space missions. Ground-based systems have been used for a number of years. The objective of the NASA program is to provide the technology to achieve a reliable space qualified cooler.

Millikelvin temperatures may be necessary for operating Weber bar-type gravitational wave detectors in space. This low temperature may provide the required sensitivity (high Q for the single crystal silicon or sapphire rod) and stability.

It is expected that the R&T program can provide the technology required to design a flight type adiabatic demagnetization cooler to achieve temperatures of a few millikelvin. The particular cooler system chosen for a given mission

will depend on the simplicity and reliability that future R&T efforts can provide for each of these systems. In all cases, cooling loads (including heat leaks) must be limited to tens of microwatts.

PRESENT/FUTURE APPLICATIONS

Table III provides a selected listing of potential future NASA missions which will use cryogenic coolers. The status of the project, the temperature required, and the potential launch date are identified for each mission identified. It should be clear to the reader that cryogenic coolers for low and ultra-low temperatures will eventually be in use on many NASA spacecraft. R&T efforts to either obtain or improve the required technology for the system design obviously have high priority. Greatly improved data with increased accuracy and resolution could be obtained with new sensors. However, these sensors must be cooled to the temperatures as indicated on Table III.

CONCLUSION

An overview of the NASA/OAST Low and Ultra-Low Temperature Cooler Research and Technology Program has been presented. Its purpose has been to acquaint the reader with the future requirements for coolers, the present status of technology, and the activities in progress to provide new technology to meet the requirements of coolers for future missions.

The areas of emphasis in a Research and Technology effort will naturally change as requirements evolve and as results from the investigations become available. Thus, this paper should be considered a status report of present activities.

REFERENCES

1. Lundholm, J. G.: NASA Low and Ultra-Low Temperature Cryogenic Cooler Systems for Space Missions. ISA Transactions, Vol. 19, No. 1, pp. 3-14.
2. Sherman, A.: Cryogenic Cooling for Spacecraft Sensors, Instruments, and Experiments. Aeronautics and Astronautics, Vol. 16, No. 11, November 1978, pp. 39-47.
3. Donohoe, M.; and Sherman, A.: Radiant Coolers-Theory, Flight Histories, Design Comparisons and Future Applications. AIAA Paper No. 75-184, Presented at 13th Aerospace Sciences Meeting, Pasadena, CA, January 1975.
4. Nast, T. C.; Barnes, C. B.; and Wedel, R. K.: Development and Orbital Operation of a Two-Stage Solid Cryogen Cooler. Journal of Spacecraft and Rockets, Vol. 15, No. 2, pp. 85-91, Mar-Apr 1978.
5. Steyart, W. A.: Rotating Carnot Cycle Magnetic Refrigerator for Use Near 2K. Journal Applied Physics, Vol. 49, No. 3, March 1978, pp. 1227-1231.
6. Urbach, A. R.; Vorreiter, J.; and Mason, P.: Design of a Superfluid Helium Dewar for the IRAS Telescope. Presented at 7th International Cryogenic Engineering Conference, July 4-7, 1978.

TABLE I.- GENERAL NASA MISSION CATEGORIES REQUIRING LOW AND ULTRA-LOW TEMPERATURE COOLERS

DISCIPLINE	APPROXIMATE TEMPERATURE RANGE (KELVIN)	APPROXIMATE RANGE OF REFRIGERATION LOAD
APPLICATIONS MISSIONS (WEATHER, EARTH RESOURCES, POLLUTION MONITORING, ETC.)	10-100	MILLIWATTS TO 10 WATTS
HIGH ENERGY AND GAMMA-RAY, ASTRONOMY	4-100	MILLIWATTS TO 10 WATTS
INFRARED ASTRONOMY	0.3-10	MICROWATTS TO 100 MILLIWATTS
RELATIVITY MISSIONS	0.001-1.5	MICROWATTS TO MILLI-WATTS
SUPERCONDUCTING DEVICES	1-15	WIDE RANGE
BASIC RESEARCH EXPERIMENTS	1-10	<100 MILLIWATTS

TABLE II.- LOW AND ULTRA-LOW TEMPERATURE SPACE COOLING TECHNIQUES

COOLING TECHNIQUE	APPROXIMATE PRACTICAL* TEMPERATURE RANGE (°K)	APPROXIMATE RANGE OF* USABLE REFRIGERATION LOAD FOR 1 YEAR MISSION
RADIANT COOLERS	70-100	0-100 MILLIWATTS
STORED SOLID CRYOGEN COOLERS	10-90	0-800 MILLIWATTS
MECHANICAL COOLERS	4-100	0-100 WATTS
ADSORPTION COOLERS	0.2-100	0-10 WATTS
MAGNETIC COOLERS	0.2-100	0-10 WATTS
STORED LIQUID HELIUM COOLERS	1.5-5.2	0-100 MILLIWATTS
He ³ COOLERS	0.3	0-100 MICROWATTS
ADIABATIC DEMAGNETIZATION COOLER	0.001-0.3	0-100 MICROWATTS

* THESE VALUES ARE THE GOALS OF THE DESIGN CONCEPTS

TABLE III.- POTENTIAL FUTURE MISSIONS REQUIRING LOW TEMPERATURE COOLERS

<u>Mission</u>	<u>Status</u>	<u>Temperature Required (°K)</u>	<u>Potential Launch Date</u>
A. <u>Infrared Astronomy Missions</u>			
1) Infrared Astronomy Satellite	Approved Project	3	1981
2) Shuttle Infrared Telescope Facility	Study & R&D	1.8	1986
3) Small Infrared Telescope	Approved-Spacelab	3	1982
4) Cosmic Background Explorer	Study	1.8	1984
5) Outer Planets Radiometry	Concept	1.8	Late 1980's
B. <u>Atmospheric Physics</u>			
1) Cryogenic Limb Scanning Interferometer Radiometer (CLIR)	Study	10	1985
2) Upper Atmosphere Research Satellite	Study	10 (Several Instruments)	1985
3) Laser Heterodyne Spectrometer	Study-Spacelab	15 to 45	1986
C. <u>Relativity Missions</u>			
1) Gyroscope Test of General Relativity	Pre-Project R&D	1.5	Late 1980's
2) Gravitational Wave Detection	Concept	0.001 to 0.1	
D. <u>Particle/Gamma Ray Astronomy Missions</u>			
1) Gamma Ray Observatory	Study	80	1984
2) Broadband X-ray Experiment	Approved	80	1985
3) Cosmic Ray Experiment	Study-shuttle	4	1986

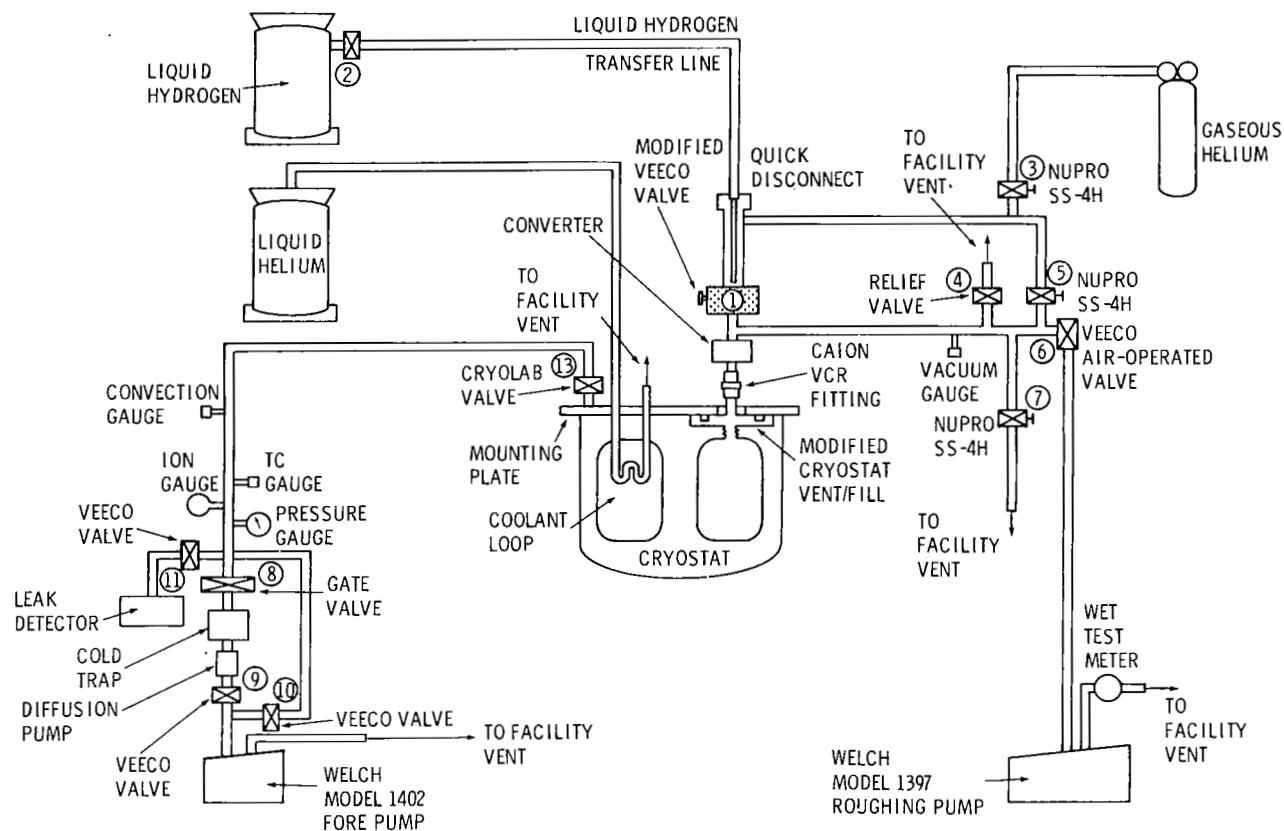


Figure 1.- Solid hydrogen test schematic.

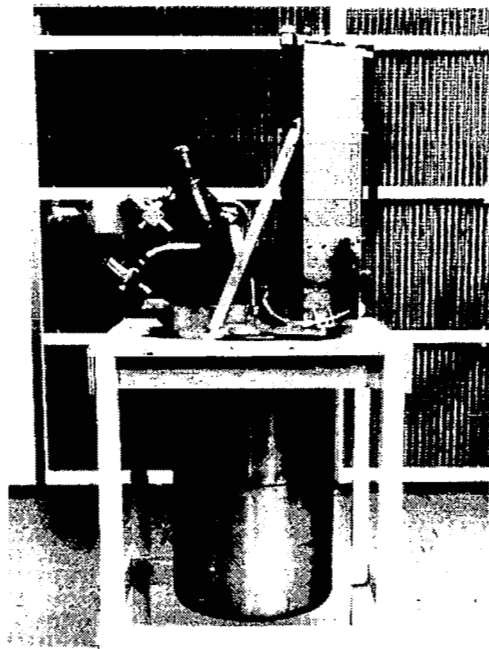


Figure 2.- Solid cryogen cooler (built by Lockheed Palo Alto) for Limb Radiance Inversion Radiometer (LRIR).

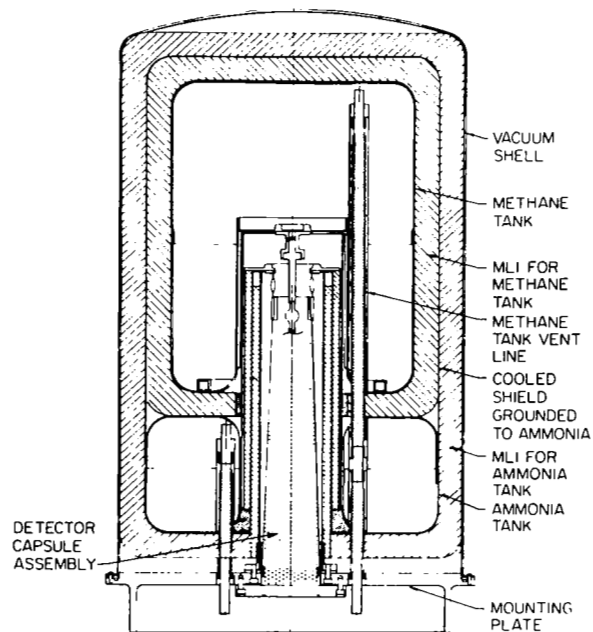


Figure 3.- Cut-away view of solid cryogen cooler (built by Lockheed Palo Alto).

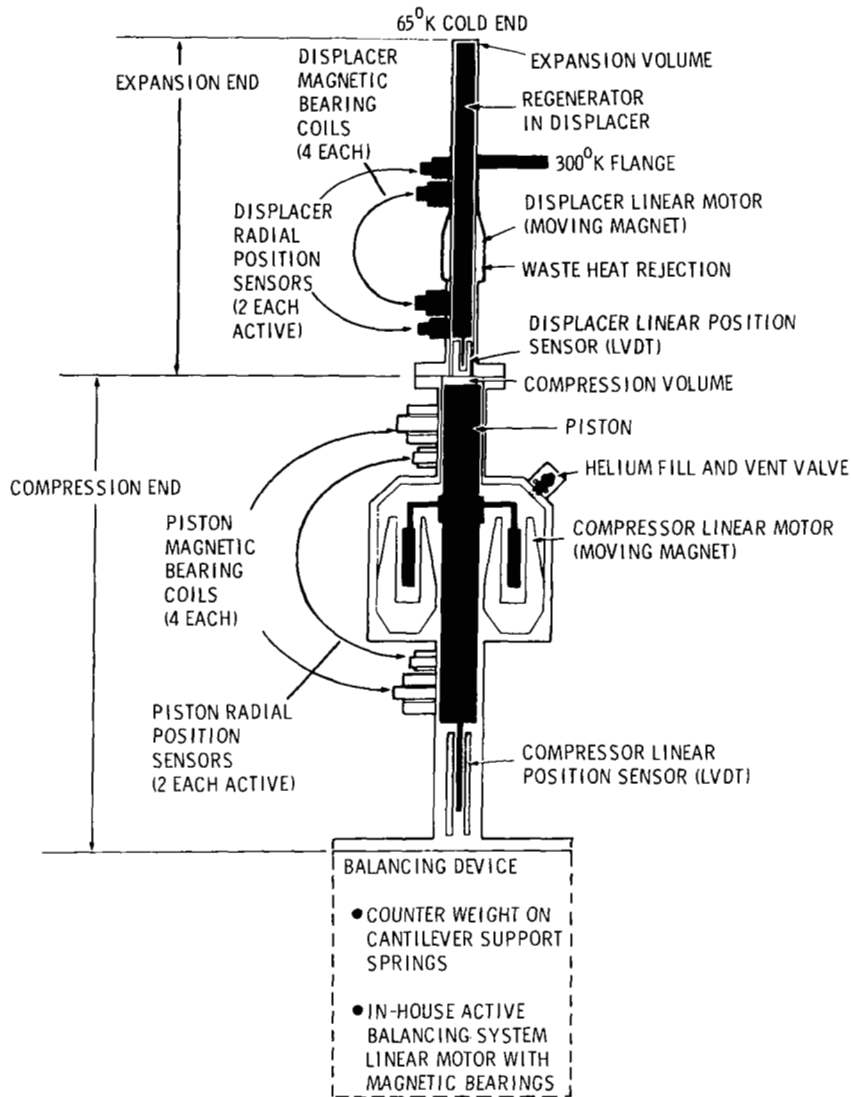


Figure 4.- Single expansion cryogenic cooler with linear magnetic suspension.

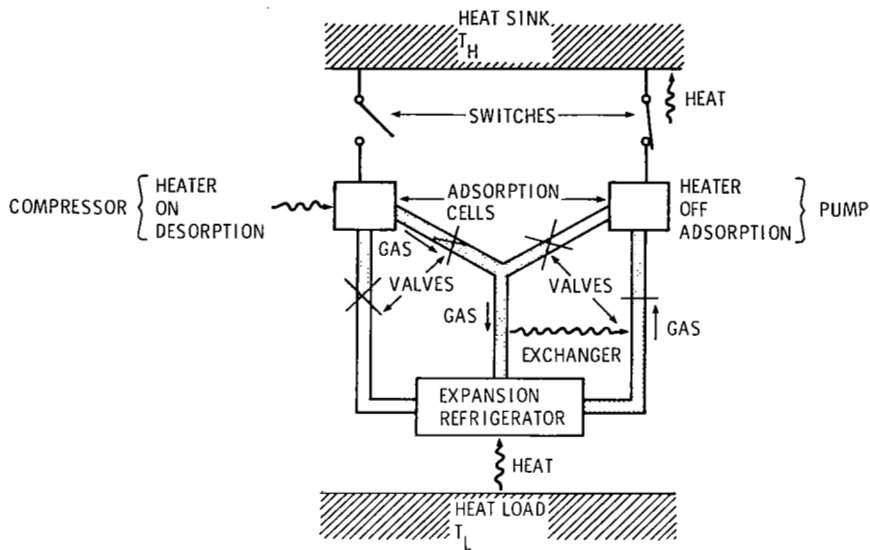


Figure 5.- Adsorption pump-compressor for expansion refrigerator.

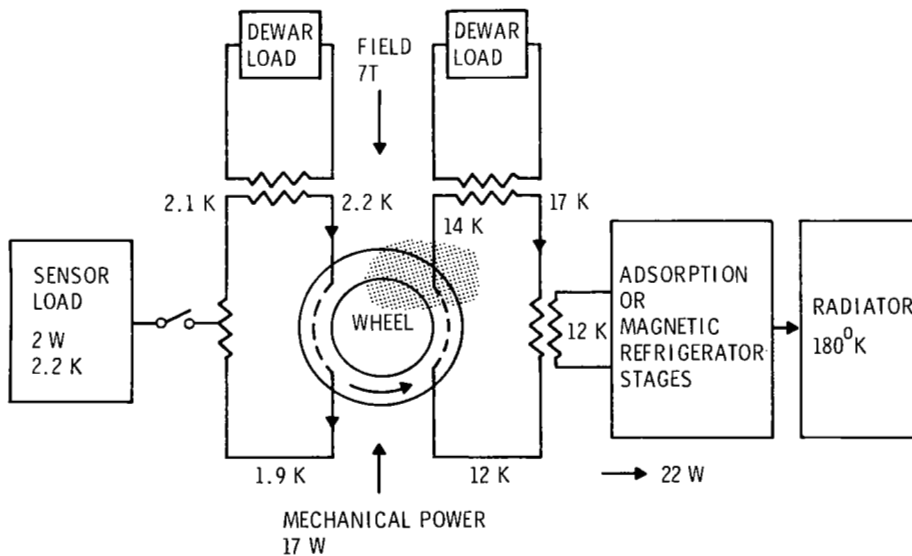


Figure 6.- Concept for spacecraft low temperature magnetic refrigerator.

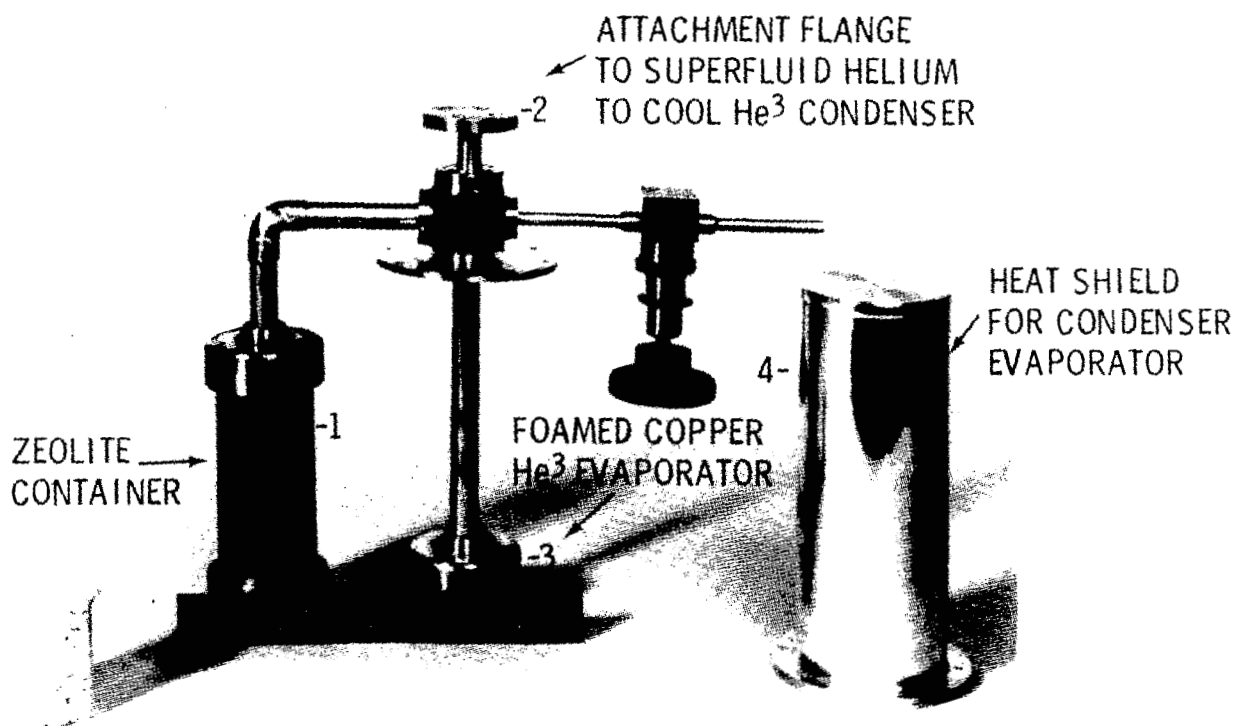


Figure 7.- He^3 cooler prototype designed by Ames Research Center.

ATMOSPHERIC STUDIES AND APPLICATIONS WITH INFRARED

HETERODYNE DETECTION TECHNIQUES

Robert T. Menzies

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91103

ABSTRACT

With the advent of stable, cw infrared lasers with moderate output power capabilities in the mid 1960's came the idea of using heterodyne detection, a mature radio frequency technique, in the infrared. This technique was first applied in radar and communications applications; however, a few years later it became apparent that infrared lasers and heterodyne detection techniques would be useful in a number of atmospheric measurement applications. The objective in this paper is not to attempt a survey of atmospheric measurement applications, many of which are represented by papers in this conference proceedings, but to indicate the level to which these techniques are being put to use in field measurements, using two developed flight instruments as examples.

INTRODUCTION

Heterodyne detection techniques are important in a variety of atmospheric measurement applications, because of the improvement in overall sensitivity which can be achieved when using high spectral resolution infrared measurements to probe the atmosphere. An infrared spectrum of the atmosphere contains a wealth of information about its chemical composition, and a careful study of the "window" regions (e.g., 3-4 μm , 8-14 μm) with a high resolution instrument will result in quantitative information about a large number of important trace constituents. The high spectral resolution usually enhances sensitivity and minimizes overlap or interference effects with other molecules. Heterodyne detection in the infrared depends on the availability of suitable lasers which can be used as local oscillators, transmitters, or both. Laser transmitters which emit at various frequencies in the "window" regions are particularly desirable, for one can use these devices to probe atmospheric regions without having to depend on sunlight or on the thermal emission properties of the species. Heterodyne radiometers which observe thermal radiation spectra are more useful at the longer wavelength "window" region, because the brightness of emitters at atmospheric temperatures peaks in this region and falls off rapidly at shorter wavelengths. Consequently, lasers which can be used as local oscillators in this wavelength region have special appeal. The grating tunable CO₂ lasers have been used for this purpose in several atmospheric measurement applications (Refs. 1-4), including studies of other planetary atmospheres which are

discussed in this Conference. The lead salt diode lasers are now being used more frequently in various measurements, and their use will certainly expand along with improvements in device technology.

In this paper, certain representative examples of atmospheric measurements with instruments employing heterodyne detection will be discussed. The examples chosen are flight instruments which have been used in several field measurements, and are indications of the potential for use of laser heterodyne technology in applications requiring ruggedized design. Several other papers in this session of the HST Conference will address recent ground-based measurements with other systems, and plans for future earth-orbiting laser heterodyne radiometers.

OZONE MEASUREMENTS WITH A LASER ABSORPTION SPECTROMETER

An airborne instrument which employs a downward-pointing laser transmitter and a matching heterodyne receiver can be used to measure pollutant concentrations in the vertical path between the instrument and ground, with modest transmitter power requirements. With a 1 watt laser and a 5 cm collecting aperture, the instrument can detect the small fraction of transmitted laser radiation which scatters off the ground and is collected at the aperture, with high enough signal-to-noise ratio to make a measurement in a few seconds integration time, at altitudes as high as 15 km. This type of instrument, measuring differential absorption at preselected wavelengths, has a high sensitivity to trace constituents at any altitude in the troposphere, whereas passive radiometric sensors are often ineffective in this application because the low altitude pollutants are at nearly the same temperature as the earth's surface.

Tropospheric ozone measurements have been made using an active, nadir-looking laser absorption spectrometer (LAS) in NASA aircraft over the past three years (Ref. 5). The LAS employs two discretely tunable CO₂ waveguide lasers and two heterodyne receivers, each tuned to detect signal returns from one of the two transmitted frequencies. A single telescope, with a 15 cm diameter primary mirror, is divided into four non-overlapping sub-apertures, two for the transmitted wavelengths and two for the independent heterodyne receivers. The application of the LAS to the mapping of ozone distributions over an urban area is shown pictorially in Figure 1. The LAS optical head, with one side cover panel removed, is shown in Figure 2. This illustrates the compactness which can be achieved with the use of waveguide lasers. In flight operation, one laser has been tuned to either the P(22) or P(24) line, which occur at frequencies where ozone absorbs very little, and the other laser has been tuned to either the P(12) or P(14) lines. Absorption coefficient data such as that shown in Figure 3 are used to calculate ozone column abundances from differential absorption measurements.

The first series of LAS flights, in the Beechcraft Queen Air February 28 through March 2, 1977, demonstrated that the measurement concept was sound. Using an integration time of a few seconds, the received signals showed very little noise. The high dynamic range capability of the receivers was sufficient

to handle large albedo fluctuations and speckle fluctuations and effectively time average them as we had hoped. The albedo fluctuations with large spatial scale (periodicities longer than 10 seconds) were present on both receiver channels and were well correlated with each other. Figure 4 is a portion of the strip chart recording taken February 28, indicating a passage from coastal land to water near Ventura, California, at a 2.5 km altitude. The high correlation between channels, as the aircraft passed over a variety of terrain, is evident. As a result, the ratio of the two receiver signals was a viable indication of differential absorption due to the atmosphere. There is very little interference from other tropospheric species at these wavelengths. However, measurements taken during recent flights over various parts of the continental U.S. indicate the possibility of interference from spectral albedo features of certain types of terrain. This effect is noticeable during low altitude flights (e.g., 800-1000 meters altitude), when it becomes necessary to measure differential absorption to a 2-3% accuracy level in order to accurately measure background tropospheric concentrations.

Recently the LAS was flown in the Southeastern Virginia area, along with a Cessna 402 which was instrumented with in-situ ozone measurement instruments, in a correlative ozone measurement program. The Cessna operation was directed by NASA Langley Research Center personnel. Data taken during Cessna spirals were compared with ozone column abundance data from the LAS. An example of these data is pictured in Figure 5. In this case, the LAS flew repeatedly over a location near Cape Charles, at three altitudes, and the results were compared with vertical spiral data. These results indicate the agreement which can be obtained when flying over water, or terrain for which spectral albedo features are either absent or previously calibrated. Various means for reducing LAS susceptibility to terrain spectral albedo are under investigation.

PASSIVE HETERODYNE RADIOMETRY

The wavelength dependence of heterodyne radiometer sensitivity to thermal radiation is shown in Figure 6. It is apparent that the sun is a good source for high spectral resolution atmospheric transmission measurements. The factor plotted in Figure 6 is multiplied by the radiometer factor, $(B_{IF}\tau)^{1/2}$ and various quantum and optical efficiency factors to obtain the signal-to-noise ratio, where B_{IF} is the receiver IF bandwidth and τ is the integration time. When the atmosphere is transparent at a frequency corresponding to a wavelength near 10 micrometers, the S/N ratio for a 20 MHz bandwidth and a 10 second integration time is about 6000. In practice, optical losses, instabilities of the solar image on the photodetector, and atmospheric effects degrade this S/N by several dB; however, there remains sufficient S/N to conduct measurements of trace species with high accuracy.

In order to remotely measure trace constituents in the stratosphere, a technique with very high sensitivity involves putting a solar heterodyne radiometer on a high altitude platform, e.g., a balloon gondola, and measuring the solar absorption spectrum from stratospheric altitudes, near sunrise or sunset. The solar occultation geometry is pictured in Figure 7. From a location in the upper stratosphere, an instrument which views the sun at a zenith angle of

slightly greater than 90° can interact over a total path length of 500 km with a portion of the stratosphere whose vertical extent is 3 km. Typical geometric weighting functions associated with this type of measurement are given in Figure 8. The high sensitivity, high vertical resolution capability, and freedom from attenuation or interference due to tropospheric or lower stratospheric species (e.g., ozone) makes this type of measurement attractive. A balloon-borne laser heterodyne radiometer (LHR) was developed at JPL for stratospheric flights beginning March, 1978, and extending through the present, in order to measure reactive trace species. Man-made halocarbons have been postulated as potentially dangerous sources of chlorine in the stratosphere (Refs. 6, 7); however, measurement data for several key species in the ozone destruction reactions, such as chlorine monoxide, are scarce. In Figure 9 are presented chlorine monoxide profiles observed during two separate LHR measurements, along with a current photochemical model prediction. (The launch location for the two balloon-borne measurements was Palestine, Texas.) Measurements indicate peak mixing ratios which are slightly higher than model predictions indicate. These data are not inconsistent with data from several in-situ measurements of chlorine monoxide by Anderson and colleagues (Ref. 8). Further measurements of chlorine monoxide, along with simultaneous measurements of other reactive stratospheric species, are planned for the near future.

REFERENCES

1. Menzies, R. T., "Laser Heterodyne Detection Techniques," in Laser Monitoring of the Atmosphere, E. D. Hinkley, editor (Springer Verlag, Berlin), 1976.
2. Mumma, M. J., T. Kostiuik, and D. Buhl, Opt. Eng. 17, 50 (1978).
3. Peyton, B. J., et. al., NASA TMX-73630, (1978).
4. Menzies, R. T., Geophys. Res. Lett. 6, 151 (1979).
5. Menzies, R. T. and M. S. Shumate, J. Geophys. Res. 83, 4039 (1978).
6. Molina, M. J. and F. S. Rowland, Nature 249, 810 (1974).
7. Hudson, R. D. and E. I. Reed, editors, The Stratosphere: Past, Present, and Future, NASA Reference Publication 1049, 1979.
8. Anderson, J. G., J. J. Margitan and D. H. Stedman, Science 198, 501 (1977).

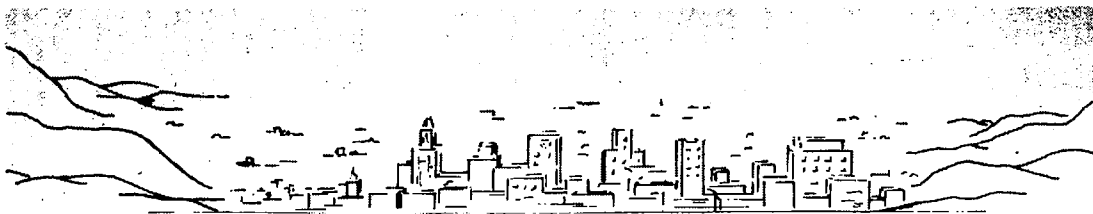
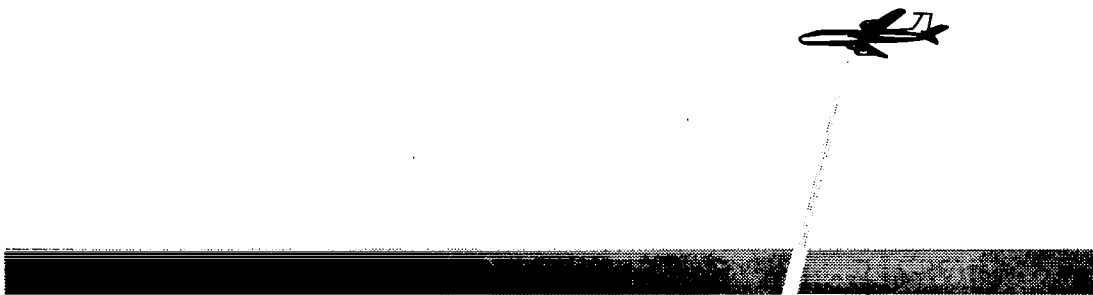


Figure 1.- The Laser Absorption Spectrometer, mounted in an aircraft, can map the distribution of ozone and other selected species in an urban area.

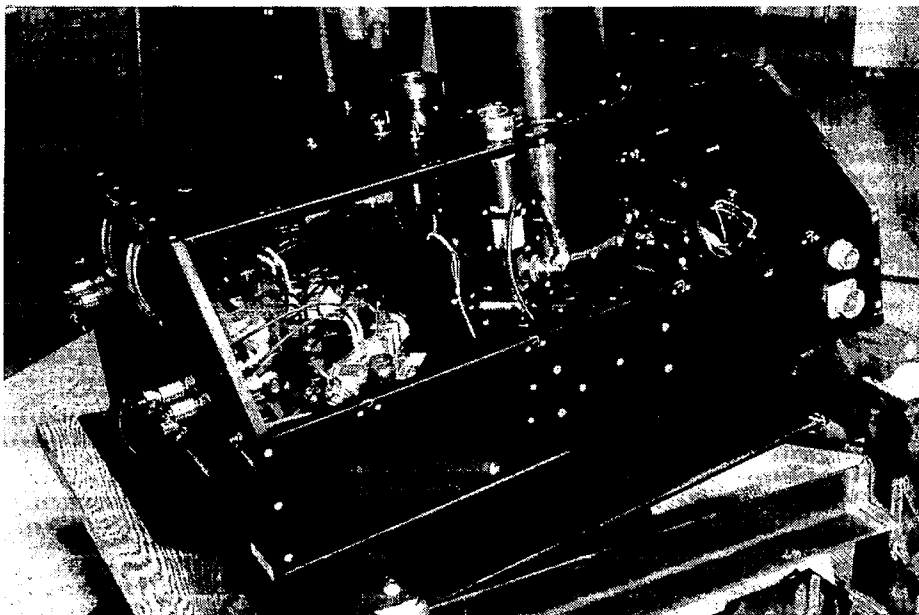


Figure 2.- LAS optical head with one panel removed. One waveguide laser, with its white BeO section visible, can be seen in the foreground near the left end. Photodetectors are mounted in liquid nitrogen dewars clamped to the telescope tube.

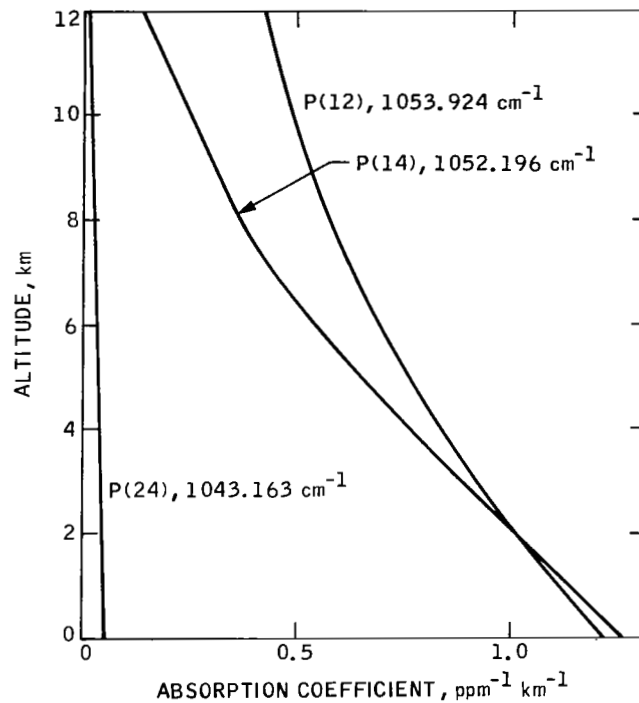


Figure 3.- Tropospheric ozone absorption at selected CO₂ laser frequencies.

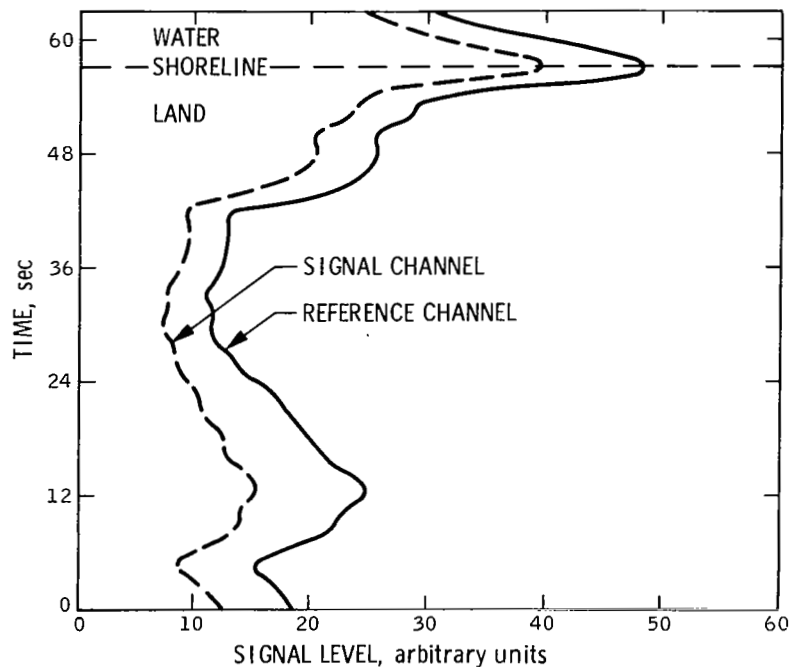


Figure 4.- Reproduction of a portion of the LAS strip chart recording, showing the high degree of correlation between the signals at the two receiver channels.

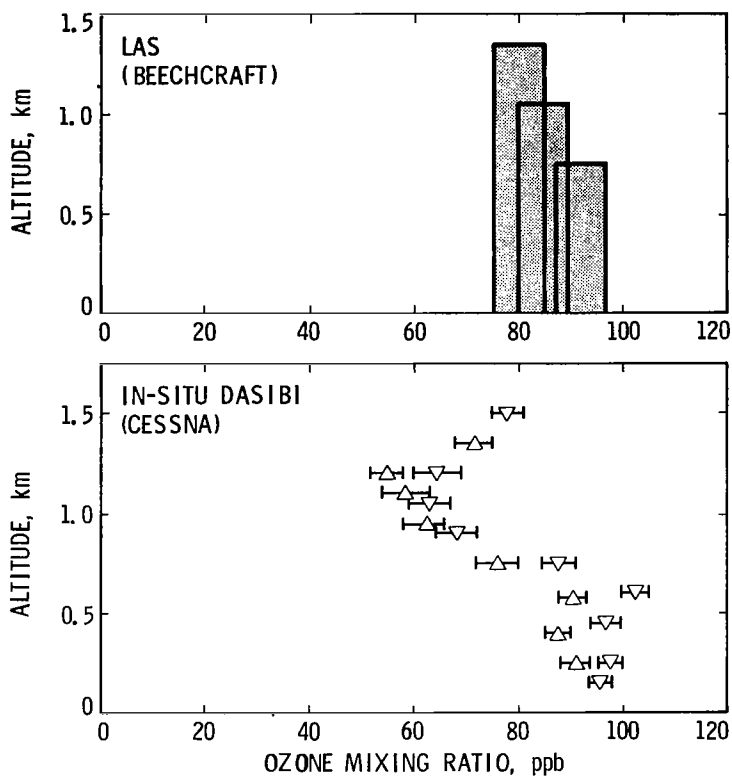


Figure 5.- Ozone measurements over Cape Charles, Virginia, 9 August 1979.

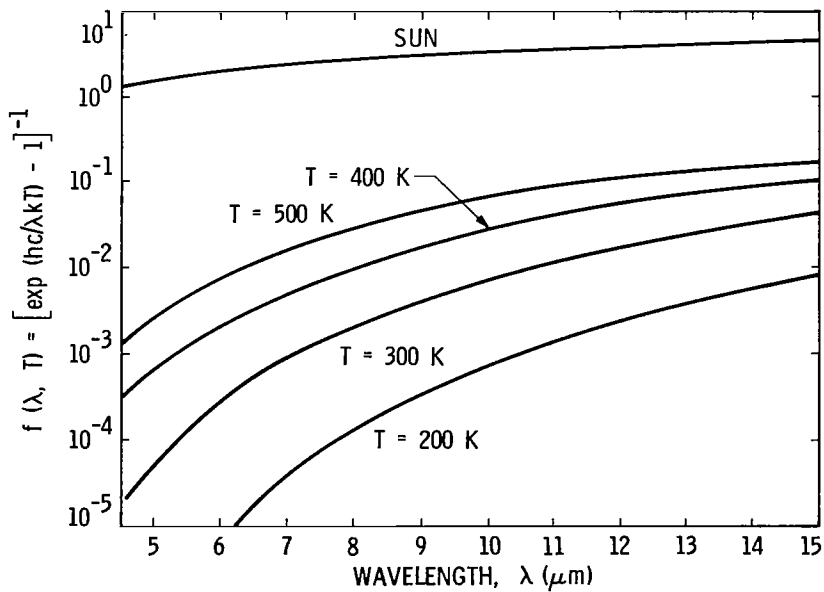


Figure 6.- Heterodyne radiometer sensitivity to thermal radiation.

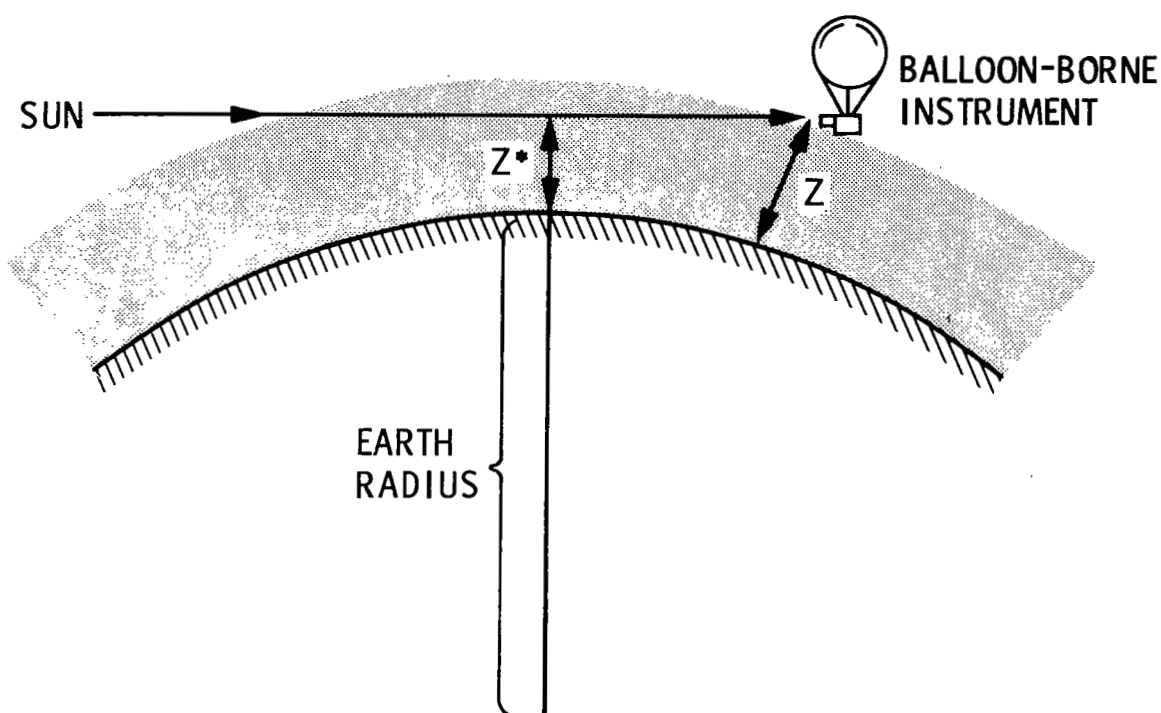


Figure 7.- Solar occultation geometry using a balloon-borne instrument.

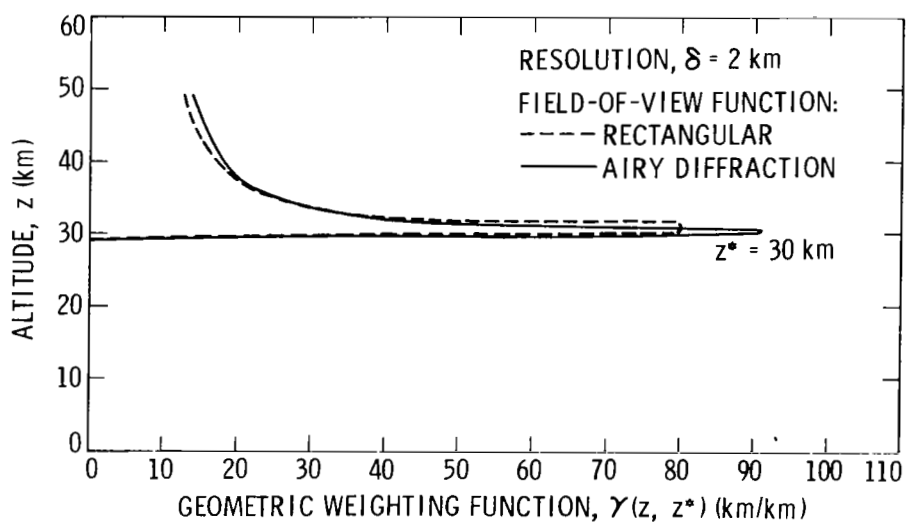


Figure 8.- Limb sensing geometric weighting functions.

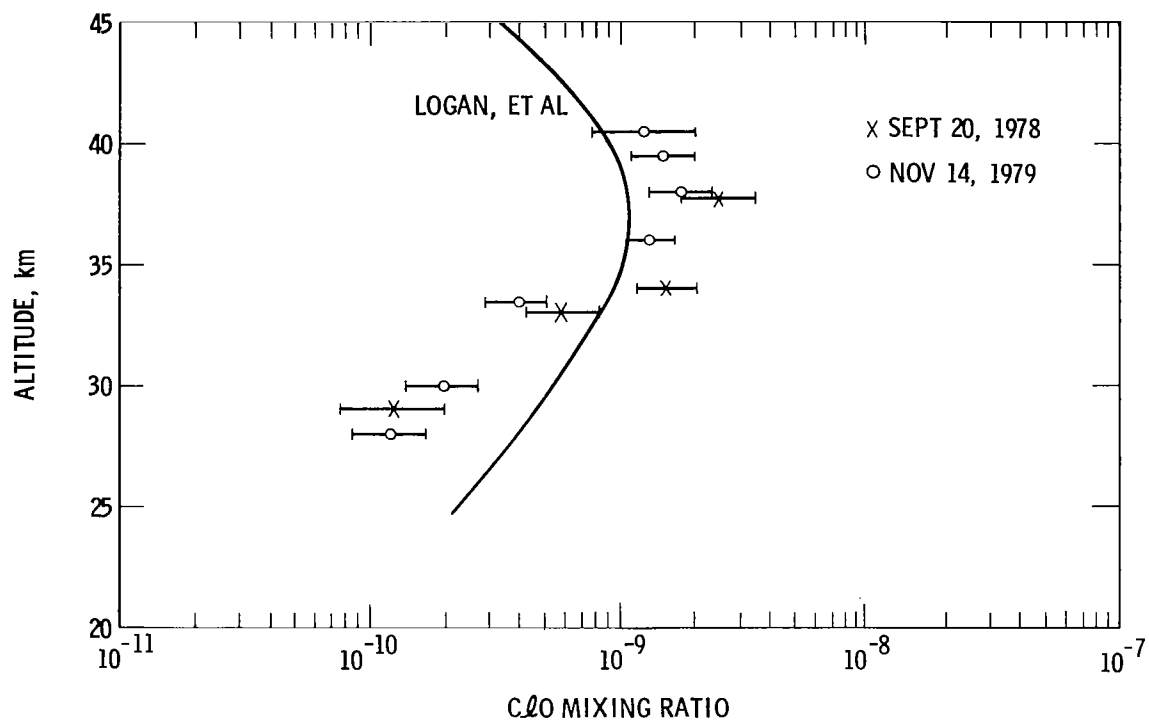


Figure 9.- LHR measurements of chlorine monoxide profiles.



INFRARED HETERODYNE RADIOMETER FOR AIRBORNE ATMOSPHERIC TRANSMITTANCE MEASUREMENTS*

J. M. Wolczok, R. A. Lange, and A. J. DiNardo
Eaton Corporation, AIL Division, Melville, NY 11747

SUMMARY

An infrared heterodyne radiometer (IHR) with an optical bandwidth of 10^{-3} cm^{-1} was used to measure atmospheric transmittance at selected hydrogen fluoride ($2.7 \text{ }\mu\text{m}$) and deuterium fluoride ($3.8 \text{ }\mu\text{m}$) laser transitions. The IHR was installed aboard a KC-135 aircraft for an airborne atmospheric measurements program that used the sun as a backlighting source for the transmission measurements. The critical components are: a wideband indium antimonide (InSb) photomixer, a CW HF/DF laser LO, a radiometric processor, and a 1900K blackbody reference source. The measured heterodyne receiver sensitivity (NEP) is $1.3 \times 10^{-19} \text{ W/Hz}$, which yields a calculated IHR temperature resolution accuracy of $\Delta T_S/T_S = 5 \times 10^{-3}$ for a source temperature of 1000K and a total transmittance of 0.5. Measured atmospheric transmittance at several wavelengths and aircraft altitudes from 9.14 km (30,000 ft) to 13.72 km (45,000 ft) were obtained during the measurements program and have been compared with values predicted by the AFGL Atmospheric Line Parameter Compilation.

INTRODUCTION

One of the critical issues in the consideration of HF and DF laser transmission is the atmospheric molecular absorption. Although molecular absorption calculations have been carried out in the 2 to 5 μm region (ref. 1), limited ground-based field measurements (ref. 2) and no high altitude measurements have been previously reported.

*Program sponsored by Air Force Weapons Laboratory, Air Force Systems Command, Kirtland AFB, New Mexico under Contract F29601-76-C-0045.

Captain Terrence F. Deaton of the Air Force Weapons Laboratory was instrumental in assembling and conducting the entire flight measurements program.

A high altitude atmospheric measurements program was carried out by the Air Force Weapons Laboratory wherein the transmittance characteristics at the lasing transitions of hydrogen fluoride (HF) and deuterium fluoride (DF) lasers have been measured. In particular, measurements were carried out using a KC-135 aircraft platform at a variety of altitudes. Atmospheric transmittance along a path from a KC-135 aircraft platform to the sun was measured using the same technique previously employed to measure the high altitude transmission at $C^{12}O_2^{16}$ laser wavelengths (ref. 3). These measurements were carried out on the $P_2(7)$, $P_2(6)$, $P_2(5)$, $P_2(4)$, $P_2(3)$, and $P_1(4)$ transitions (2.8706 to 2.6401 μm) of the HF laser and the $P_3(9)$, $P_3(8)$, $P_3(7)$, $P_3(6)$, $P_2(9)$, $P_2(8)$, $P_2(7)$, $P_2(6)$, and $P_2(5)$ transitions (3.9654 to 3.698 μm) of the DF laser. The measurement results indicate a unity atmospheric transmittance on all of the DF laser transitions, and a generally lower transmittance than predicted for the HF laser transitions.

The IHR measured atmospheric transmittance within a spectral bandwidth of 10^{-3} cm^{-1} about all of the HF and DF laser transitions. The measured receiver sensitivity (NEP) of the IHR is $1.3 \times 10^{-19} \text{ W/Hz}$, which yields a calculated IHR temperature resolution accuracy of $\Delta T_S/T_S = 5 \times 10^{-3}$ for source temperatures, $T_S = 1000\text{K}$ and transmittance $\alpha = 0.5$. Measurement accuracies near 0.5 percent have been calculated at HF and DF laser wavelengths.

IHR OPTICAL TRAIN

The IHR optical train, shown in figure 1, consists of an objective lens that brings incident solar or calibration blackbody radiation to a focus, where it is Dicke switched (alternately referenced to an ambient temperature blackbody) by a motor-driven chopper wheel. The chopper wheel, which serves as the ambient blackbody, is coated with high-emissivity paint and held at a fixed reference temperature. The radiance passing through the chopper, and that emitted by the chopper blade, is collected and focused onto the PV:InSb photomixer by a relay lens. A plane-parallel beam from the local oscillator laser is directed onto the photomixer by a small injection mirror located directly behind the relay lens.

The optical Dicke switch is located at the focal point of the objective lens in order to provide square wave chopping of the sources, with a duty cycle of 50 percent. The position of the Dicke switch is electronically monitored by chopping an auxiliary light source to provide a synchronizing signal for the radiometric processor. This sync signal is sent to the electronics console, in the form of synchronizing pulses, so that detection can be synchronous with the Dicke switch position.

Provisions were made for temperature calibration of the IHR output with respect to a 300K and a 1900K blackbody calibration source. Since the unattenuated solar temperature is approximately 5750K (ref. 4), the use of a 1900K calibration blackbody (rather than 1300K) provided improved IHR measurement accuracy at temperatures greater than 1300K due to reduced extrapolation of the calibration signals. The IHR internal calibration sources can be inserted by using a flip mirror ahead of the objective lens.

The photomixer responds to the incident radiation energy at infrared frequencies above and below the local oscillator frequency and is band-limited by IF filters. This detected signal is amplified and processed (ref. 5). The frequency resolution is fixed at twice the IF bandwidth and, therefore, the IHR can provide an extremely high spectral resolving capability. Indeed, the present instrument has provision for a 50-MHz IF filter that yields an infrared spectral resolution of $3 \times 10^{-3} \text{ cm}^{-1}$ at $\lambda = 2.7 \text{ }\mu\text{m}$. The necessity for such high spatial resolution is evident when one considers the extremely active and discrete line spectra of atmospheric species in the 2 to 5 μm region. For the intended measurement, the IHR instrument must have a spectral resolution that is comparable with the laser line resolution in order to measure a meaningful atmospheric transmittance.

The IHR optical system is shown in figure 2. The calibration insertion mirror can be seen adjacent to the objective lens. This mirror, together with the 300K calibration blackbody, can be positioned remotely from the electronics control console. The unit is RFI shielded, and all power leads enter the optical package via an RFI filter box. These precautions have been taken to ensure minimum interference from the aircraft environment and the HF (DF) laser LO, which is potentially a major source of RF interference.

Solar radiation is brought into the radiometer via an optical tracking system. The tracker is a video system with 30-cm (12-in.) diameter primary and secondary mirrors. The target acquisition is manual, and the control is then handed over to an on-board computer. Separate optical paths are used for: (1) the visible light to the video camera, and (2) the infrared radiation to the IHR. The "bicolor" window consists of a 15 cm (6 in.) x 25 cm (10 in.) quartz window for the visible radiation, and a 10 cm (4 in.) x 25 cm (10 in.) germanium window for the infrared radiation. The unobscured field of view of the tracker is approximately $\pm 6^\circ$ in azimuth and 0° to 45° in elevation.

IHR ELECTRONICS

Since the IHR uses an ambient temperature reference (the Dicke switch chopper wheel), direct null-balanced radiometer operation is not possible when observing hot sources such as the sun. The method employed to deal with these hot sources is the utilization of gain modulation in the radiometer (ref. 5). The gain modulation circuitry alters the system gain during alternate phases of

the Dicke switch cycle, thereby adjusting that portion of the detected signal that represents the reference thermal energy. This enables null-balance operation of the IHR for source temperatures that are much larger than the reference temperature. Gain modulation is achieved by electronically switching a fixed attenuator into the null-balancing loop of the synchronous detector at the Dicke switch rate. The IHR temperature resolution with gain modulation is identical to that obtained when the IHR is null balanced by means of a reference temperature that is equal to the source temperature. Therefore, no degradation in IHR performance is experienced due to any variations in laser LO power or IHR gain which occur slowly compared to the Dicke switch rate.

The IF output from the InSb photomixer is amplified and processed in an AIL 777 Radiometer Processor. The processor provides synchronous detection that employs the self-balancing gain modulation technique (ref. 3). A simplified block diagram of the processor is shown in figure 3. The output of the InSb photomixer is IF amplified, detected, and decommutated. The output level obtained while viewing the chopper wheel reference temperature is used to maintain a fixed system gain by means of feedback to the AGC amplifier located ahead of the square-law detector. The decommutated output is used to provide separate control of the processor gains, when viewing either the reference chopper wheel or the external source/calibration blackbody. The output of the synchronous detector is driven to a null balance by means of a variable digital attenuator which controls the video gain in the processor when viewing the reference source.

Temperature range and resolution control of the processor output are available to maximize the resolution of the digital output for any range of expected source temperatures. The digital attenuator output, which is directly proportional to input source radiance, has 1024 discrete steps.

IHR PHOTOMIXER AND LASER LO

A photovoltaic InSb detector with an active area of $1.5 \times 10^{-4} \text{ cm}^2$ is used as the IHR photomixer. The reverse shunt resistance of the InSb detector is of the order of 3 megohms, resulting in low values of photomixer dark current (I_D), and the requirement for an IF preamplifier that operates from a relatively high source impedance. Frequency response measurements were carried out with the InSb photomixer at 77K in the IHR configuration. The measured 3-dB cutoff frequency was $f_c = 50 \text{ MHz}$ for a reverse bias voltage of $V_B = -600 \text{ mV}$, and increased to $f_c = 60 \text{ MHz}$ for $V_B = -800 \text{ mV}$. The laser LO induced photocurrent was about 0.15 mA compared to a dark current of approximately 0.02 mA.

The CW HF (DF) laser is of the basic design which was initially reported by Hinchin (ref. 6), with appropriate modifications for airborne use. The

grating-tunable, flowing-gas, water-cooled laser has an optical cavity which is 30 cm long and provides a TEM₀₀ transverse mode. A stabilization loop is employed to continuously adjust the cavity length and maintain the laser tuned on the center of the selected transition.

IHR TEMPERATURE RESOLUTION ACCURACY

The IHR temperature resolution, T_s , is defined as the maximum change in source temperature for which no change in output signal-to-noise ratio is observable. The IHR temperature resolution is given (ref. 4) by

$$\Delta T_s = \frac{k^2 I_s^2}{h^2 \nu^2} \cdot \frac{[\exp(h\nu/kT_s) - 1]^2}{\alpha \exp(h\nu/kT_s)} \cdot \sqrt{\frac{K}{B\tau}} \left[\frac{2h\nu\alpha}{k [\exp(h\nu/kT_s) - 1]} + \int_{f_1}^{f_2} \frac{NEP(f) df}{kB} \right] \quad (1)$$

where α is the transmittance of the media and optics between thermal source and infrared photomixer, K is the sensitivity constant which is 2 for a Dicke-type receiver, T_s is the source temperature, B is the predetection (IF) bandwidth, τ is the postdetection integration time for a pure integration, ν is the infrared frequency, k is Boltzmann's constant, h is Planck's constant, and f is the IF frequency.

The receiver sensitivity (NEP) for a heterodyne receiver with a photovoltaic photomixer is given (ref. 7) by

$$\left. \begin{aligned} NEP &= \int_{f_1}^{f_2} NEP(f) df \\ NEP(f) &= \frac{h\nu}{\eta} \left\{ 1 + 2K \frac{(T_m + T'_{IF})}{q I_o} G_D [1 + (f/f_c)^2] \right\} = \frac{h\nu}{\eta'} \end{aligned} \right\} \quad (2)$$

where η is the photomixer quantum efficiency, T_m is the photomixer temperature, T'_{IF} is the effective input noise temperature of the IF amplifier, G_D is the reverse shunt conductance of the photomixer, I_o is the LO induced photocurrent, f_c is the 3-dB cutoff frequency of the photomixer, and η' is the effective receiver quantum efficiency.

The temperature resolution of the IHR is affected by optic losses, spherical aberration, coma, astigmatism, lens emissivity, lens temperature, amplitude fluctuations that occur faster than the Dicke switch rate, and the obscuration due to the LO injection mirror. The calculated resolution accuracy, $\Delta T_s/T_s$, of a gained-modulated Dicke-switched IHR is given in figure 4 for $\lambda = 3 \mu\text{m}$, $\beta\tau = 5 \times 10^9$, and a heterodyne receiver NEP of 1.3×10^{-19} W/Hz with the transmittance α as a parameter. For a source temperature of 1000°C and $\alpha = 10$ percent, the calculated temperature resolution accuracy is $\Delta T_s/T_s \approx 1.6 \times 10^{-2}$, while for $\alpha = 100$ percent, the resolution accuracy improves to $\Delta T_s/T_s \approx 1.6 \times 10^{-3}$. For the solar viewing case ($T_s \approx 5750\text{K}$ for 100 percent atmospheric transmittance), the calculated temperature resolution accuracy is $\Delta T_s/T_s = 9 \times 10^{-5}$ for a transmittance of $\alpha = 0.25$.

IHR TEMPERATURE RESOLUTION

The solar radiation, $P(v)$, collected at the IHR is the product of the incident unattenuated solar radiation, P_o , and an exponential attenuation term due to the intervening atmospheric losses. Therefore

$$P(v) = P_o \exp [-k_\alpha(v) \text{ nsec } \theta] \quad (3)$$

where $k_\alpha(v)$ is the atmospheric attenuation coefficient at selected laser LO frequency, h is the height of atmosphere, θ is the zenith angle of the sun, and $\text{nsec } \theta$ is the path length of sunlight through the atmosphere.

The IHR output voltage is proportional to the collected source energy. The slope of the logarithm of $P(v)$ plotted against $\sec \theta$ is $-k_\alpha(v) h$, and $\exp [-k_\alpha(v)h]$ is the total vertical path transmittance α (ref. 2, 4, 8, and 9). The vertical path atmospheric transmittance can therefore be experimentally determined by measuring the solar radiance while tracking the sun as the earth

turns. The measurements can be carried out at different laser local oscillator frequencies, solar angles, and altitudes.

When observing a thermal source that fills the IHR field of view, the postdetection signal-to-noise ratio is given (ref. 4) by

$$\text{SNR} = \frac{\eta' \nu (B\tau)^{1/2}}{e^{h\nu/kT_s} - 1} \quad (4)$$

Equation (4) can be rewritten as

$$\alpha = \gamma \frac{S}{N} \quad (5)$$

where γ is a constant for a fixed set of measurement parameters. Differentiating α with respect to S yields

$$d\alpha = \frac{\gamma}{N} ds \quad (6)$$

The minimum detectable signal change occurs when $ds = N$. Therefore, the accuracy of the measurement of atmospheric transmittance is given by

$$\frac{d\alpha}{\alpha} = \frac{1}{\text{SNR}} \quad (7)$$

For a typical solar observation, where $T_s = 5000\text{K}$, $B = 50\text{ MHz}$, $\tau = 10\text{ s}$, $B = 5 \times 10^9$, $\eta' = 0.2$, and a total atmospheric transmittance of $\alpha = 0.5$ from sea level to 30 km during midlatitude summer at $\lambda = 3\text{ }\mu\text{m}$, the SNR of an IHR observing the solar disc at zenith is calculated to be $\text{SNR} = 2000$, and the expected accuracy of the measurement is $d\alpha/\alpha = 0.05\text{ percent}$.

Calculations have also been carried out to determine the obtainable IHR power resolution. The resolution is a function of the effective system quantum efficiency, the calibration blackbody temperature, and the $B\tau$ product of the IHR. The smallest calculated change in irradiance that can be measured for a given input radiance is given in figure 5. It should be noted that the overall system quantum efficiency, η' , depends on both the quantum efficiency of the infrared photomixer and the heterodyne mixing efficiency which is determined by the IHR optical configuration and the photomixer size.

The PV:InSb photomixer used in the IHR has a quantum efficiency of approximately 75 percent. The IHR system loss factor is approximately 3 dB due

to both optics losses and mixing inefficiency. This results in an overall IHR system quantum efficiency of $\eta' = 37$ percent. As seen from figure 5, a 1900K blackbody source will yield a resolution of about 3 parts per 1000 as compared with a resolution of about 10 parts per 1000 for a 1300K source temperature for $\lambda = 2.9 \mu\text{m}$ and $B\tau = 10^9$.

CALCULATED AND MEASURED ATMOSPHERIC TRANSMITTANCE DATA

Calculations and measurements have been carried out to determine the expected atmospheric transmittance at various altitudes for a number of selected HF and DF laser transitions. As an example, the calculated transmittance for a midlatitude winter model (with no aerosols present) is given in table I for the $P_2(7)$ transition ($\lambda = 2.8706 \mu\text{m}$) of the HF laser (ref. 1). The transmittance values have been calculated for the optical path between the edge of the atmosphere and the selected aircraft altitude, with the solar viewing angle as a variable parameter. As can be seen in table I, the calculated transmittance for the $P_2(7)$ transition of HF increases with increasing aircraft altitude.

The IHR flight measurements program occurred between July, 1977 and January, 1978. During the flight tests, atmospheric pressure, temperature, and humidity data were obtained from point analyses supplied by the USAF Environmental Technical Applications Center (ETAC). ETAC has developed a linear regression model (ref. 10) which estimates water vapor levels above the -40° level, at which standard operational rawinsondes fail. The ETAC model is based on a series of actual soundings by the Naval Research Laboratory. Each point analysis is based on the data from a particular rawinsonde launch. For early morning flights over New Mexico, the 1200Z (5 A.M. MST) Albuquerque rawinsonde data and corresponding point analysis data were used to compute the transmittance. In a few instances, soundings from El Paso and Midland, Texas, and Grand Junction, Colorado were also used. A sample profile showing the comparison between the point analysis and the rawinsonde data of the water vapor profile is shown in figure 6. As can be seen, the agreement is excellent and the water vapor pressure at 30,000 feet is about two orders of magnitude lower than the water vapor pressure between ground level and 1.83 km (6,000 ft).

The calculated transmittance values do not account for aerosol scattering or absorption, which might have a small effect on the transmittance at the altitudes of interest. The calculated water vapor absorption, even for relatively high humidity conditions, is almost negligible for altitudes above 12.19 km (40,000 ft) for all the laser transitions which were measured. The molecular absorption is due chiefly to CO_2 , with significant N_2O absorption only on the $P_2(7)$ transition of the HF laser. The agreement of the measurements with the calculated transmittance values is generally good. Because of noise problems, due to poor HF laser stability, the IHR signal-to-noise ratio

for unattenuated solar radiation was about 30:1 and, therefore, it was difficult to measure transmission of less than $\alpha = 0.10$.

Measured atmospheric transmittance data for several selected HF laser transitions are given in figures 7 and 8. Tables II, III and IV summarize some of the measured atmospheric transmittance data for the $P_2(6)$, $P_2(7)$, and $P_1(4)$ transitions of the HF laser.

CONCLUSION

An infrared heterodyne radiometer has been designed, developed, and fabricated for use on a KC-135 aircraft. The extremely narrow bandwidth (10^{-3} cm^{-1}) IHR has been designed to operate at the HF/DF laser transitions and exhibits the following characteristics:

Wavelength	2.7 to 3.8 μm
Receiver type	Gain modulated, Dicke-switched, heterodyne
Receiver field of view	$\approx 2 \text{ mrad}$
Receiver sensitivity (NEP)	$1.3 \times 10^{-19} \text{ W/Hz}$
Dicke switch rate	500 to 2000 Hz, selectable
IF bandwidth	50 or 100 MHz, selectable
Integration time	0.1, 0.5, 1.3, or 10 s selectable
Reference temperature	Ambient
Calibration temperature	1000 to 1900K, ambient
Blackbody calibration accuracy	$\approx 1 \text{ percent}$
Blackbody calibration	$\approx 0.25 \text{ percent}$, long term $\approx 0.05 \text{ percent}$, short term
Overall measurement accuracy	0.5 to 0.1 percent

During the design and development of the IHR, consideration was also given to operation at other infrared wavelengths. Although the IHR is presently being used for solar and atmospheric transmittance measurements, there are a number of other potential applications for this unique instrument. The

infrared spectrum is rich in absorption (emission) phenomena in the 2 to 12 μm spectral region. The narrow bandwidth IHR offers the potential of spectral overlap with selected signature lines to permit the remote detection and monitoring of various atmospheric constituents. Previous atmospheric measurements near $\lambda = 10 \mu\text{m}$ (ref. 8) have resulted in the remote determination of the concentration profile of atmospheric ammonia and ozone.

TABLE I. CALCULATED ATMOSPHERIC TRANSMITTANCE FOR $P_2(7)$ TRANSITION
OF HF LASER FOR MIDLATITUDE WINTER (NO AEROSOL) AND
 θ = SUN'S ANGLE FROM ZENITH

<u>Altitude</u> <u>(km)</u>	<u>Transmittance</u>			
	<u>$\theta = 80^\circ$</u>	<u>$\theta = 70^\circ$</u>	<u>$\theta = 60^\circ$</u>	<u>$\theta = 50^\circ$</u>
10	0.9552	0.9770	0.9842	0.9877
9	0.9348	0.9664	0.9769	0.9820
8	0.8984	0.9471	0.9635	0.9715
7	0.8241	0.9064	0.9350	0.9491
6	0.6791	0.8216	0.8742	0.9007

TABLE II. ZENITH TRANSMITTANCE FOR $P_2(6)$ TRANSITION OF HF LASER
($\lambda = 2.8318 \mu\text{m}$)

<u>Altitude</u>	<u>Measured transmittance</u>	<u>Calculated transmittance</u>
9.45 km (31,000 ft)	0.924	0.918
10.67 km (35,000 ft)	0.954	0.939
11.28 km (37,000 ft)	0.987	0.968
11.89 km (39,000 ft)	0.970	0.954
12.50 km (41,000 ft)	0.982	0.964
13.72 km (45,000 ft)	1.00	0.974

TABLE III. ZENITH TRANSMITTANCE FOR $P_2(7)$ TRANSITION OF HF LASER
($\lambda = 2.8706 \mu\text{m}$)

<u>Altitude</u>	<u>Measured transmittance</u>	<u>Calculated transmittance</u>
10.67 km (35,000 ft)	0.993	0.995
11.28 km (37,000 ft)	1.0	0.993
11.89 km (39,000 ft)	1.0	0.996
12.50 km (41,000 ft)	1.0	0.998
13.72 km (45,000 ft)	1.0	0.998

TABLE IV. ZENITH TRANSMITTANCE FOR $P_1(4)$ TRANSITION OF HF LASER
($\lambda = 2.6401 \mu\text{m}$)

<u>Altitude</u>	<u>Measured transmittance</u>	<u>Calculated transmittance</u>
10.97 km (36,000 ft)	0.55	0.975
12.50 km (41,000 ft)	0.74	0.99
13.72 km (45,000 ft)	0.85	0.995

REFERENCES

1. McClatchey, R. A.; and Selby, J. E.: Atmospheric Attenuation of HF and DF Laser Radiation. Air Force Cambridge Research Lab., Report No. AFCRL-72-0312, May 1972.
2. Snemabukuro, F. I. et al: Atmospheric Transmission Measurements at HF and DF Laser Wavelengths. Appl. Opt. vol. 15, no. 5, May 1976, pp. 1115-1117.
3. Peyton, B. J.; Hoell, J. M.; Lange, R. A.; Savage, M. G.; and Allario, F.: Remote Airborne Measurements of Stratospheric Ozone. Proc. of Laser '79 Conference, Orlando, Fla. 1979.
4. Murcray, F. R.; Murcray, D. G.; and Williams, W. J.: The Spectral Radiance of the Sun from 4 μ m to 5 μ m. Appl. Optics, vol. 3, no. 12, Dec. 1964, pp. 1373-1377.
5. Peyton, B. J. et al: An Infrared Heterodyne Radiometer for High-Resolution Measurements of Solar Radiation and Atmospheric Transmission. IEEE J. Quant. Elec., QE-11, Aug. 1975.
6. Hinchey, J. J.: Operation of a Small Single-Mode Stable CW Hydrogen Fluoride Laser, J. App. Phys., no. 45, 1974, p. 1818.
7. Peyton, B. J. et al: High Sensitivity Receiver for Infrared Laser Communication. IEEE J. Quant. Electr., QE-8, Feb. 1972, pp. 252-263.
8. Peyton, B. J. et al: Infrared Heterodyne Spectrometer Measurements of the Vertical Profile of Tropospheric Ammonia and Ozone. AIAA Conference, Los Angeles, Ca., Jan. 1977.
9. King, S. R.; Hodges, D. T.; Hartwick, T. S.; and Baker, D. H.: High Resolution Atmospheric-Transmission Measurements Using a Laser Heterodyne Radiometer, Appl. Opt., vol. 112, June 1973, p. 1106.
0. Mendenhall, L. D.; Stanton, T. E.; and Henderson, H. W.: A Model for Describing the Atmospheric Water Vapor Profile above the -40°C Temperature Level. USAF Environmental Technical Applications Center Report 7584, Aug. 20, 1975.

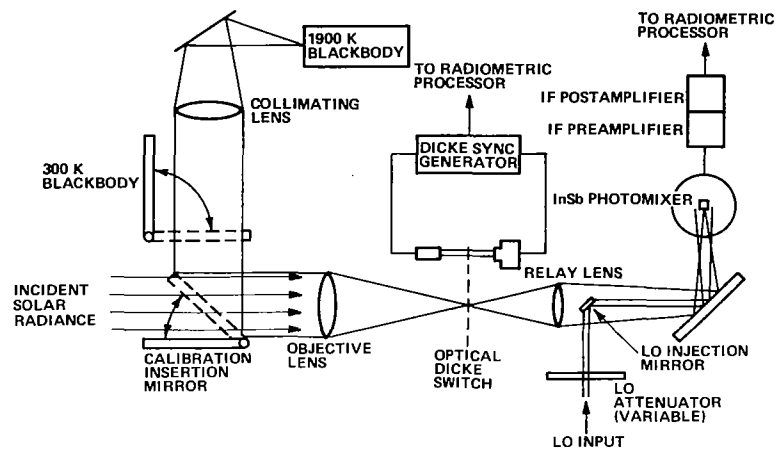


Figure 1.- Diagram of IHR optical train.

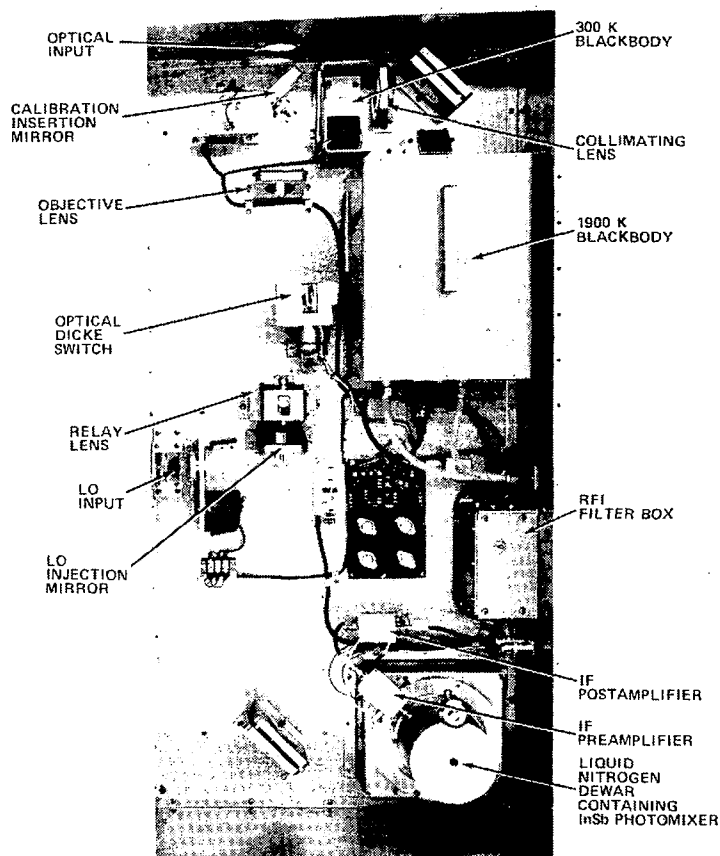


Figure 2.- Top view of IHR optical system.

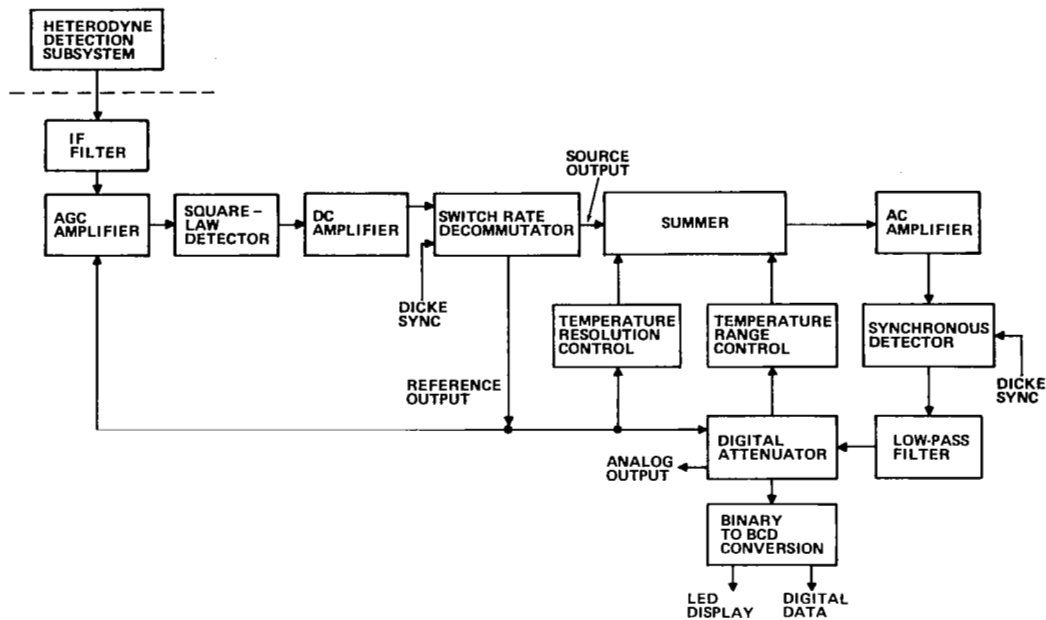


Figure 3.- Block diagram of IHR processor with self balancing gain modulation.

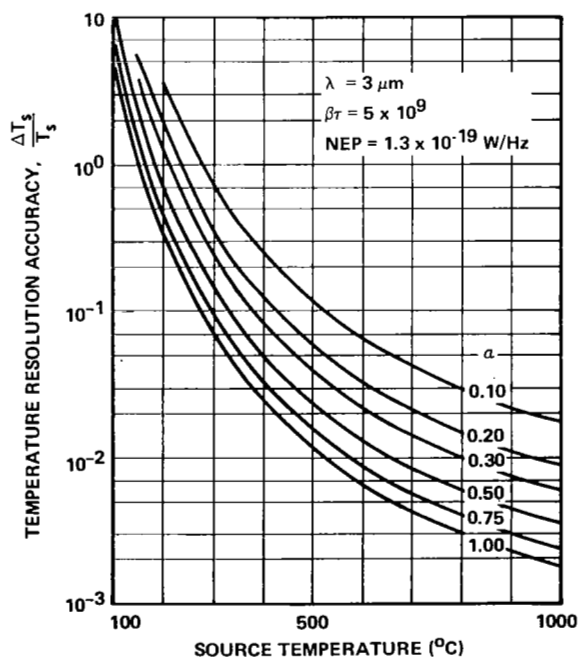


Figure 4.- Calculated IHR temperature resolution.

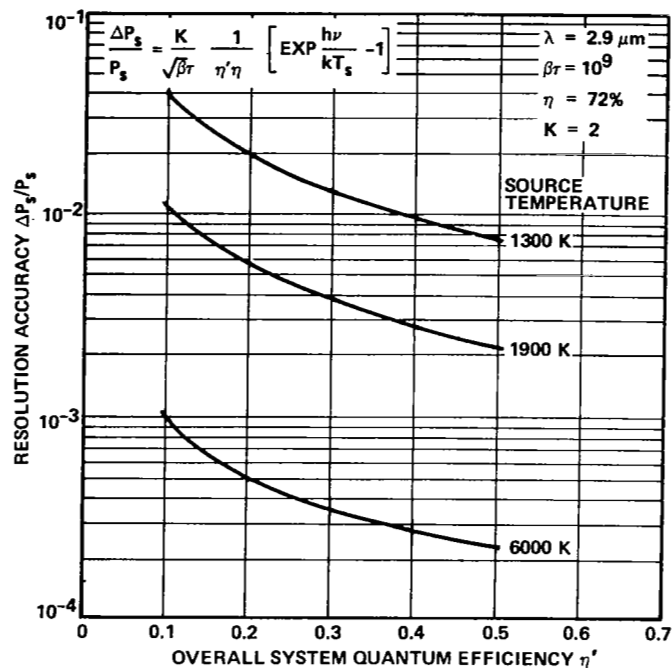


Figure 5.- Calculated IHR resolution accuracy.

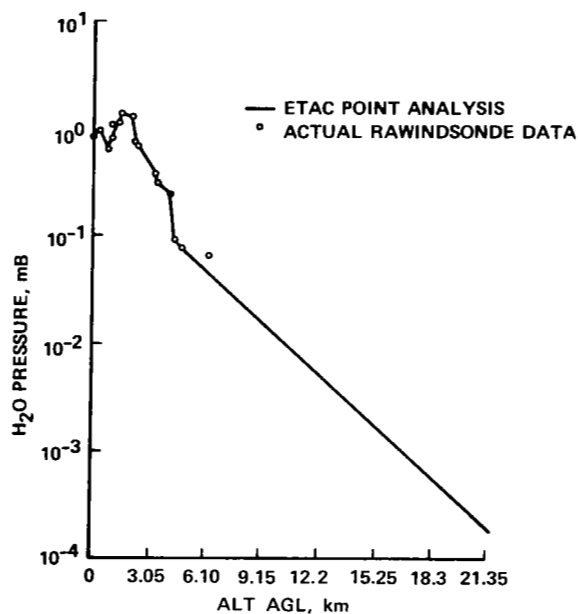


Figure 6.- ETAC point analysis water vapor profile compared to actual rawinsonde data for sounding of Bismarck, ND 1700 MST, 7 Jan. 1978.

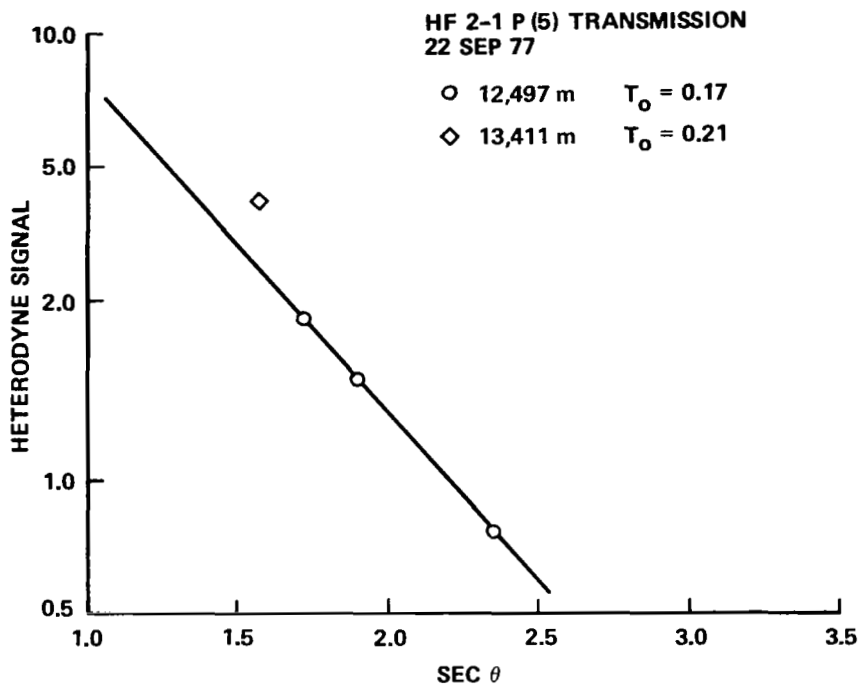


Figure 7.- HF $P_2(5)$ transmission plot for altitudes shown over Rapid City, SD, 22 Sept. 1977.

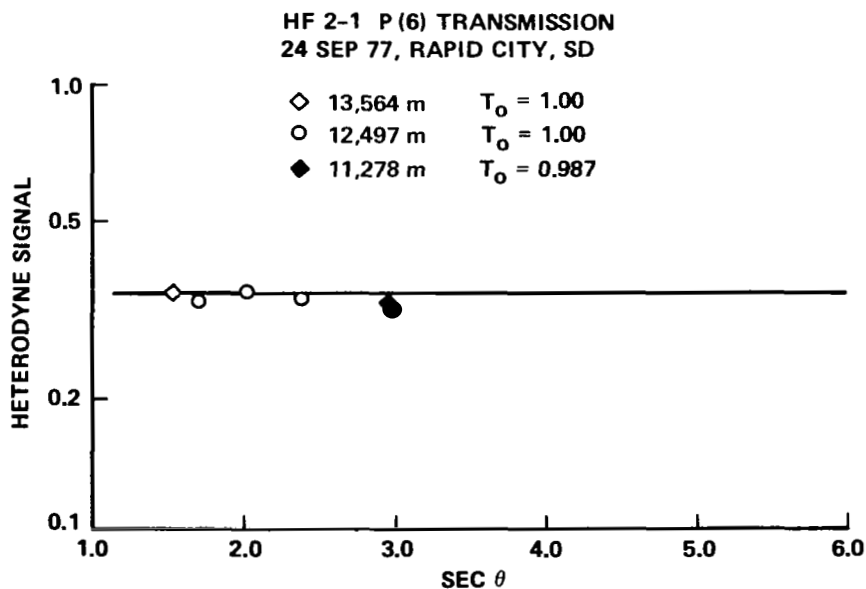


Figure 8.- HF $P_2(6)$ transmission plot for altitudes shown over Rapid City, SD, 24 Sept. 1977.



DEVELOPMENT AND PERFORMANCE OF A LASER HETERODYNE
SPECTROMETER USING TUNEABLE SEMICONDUCTOR
LASERS AS LOCAL OSCILLATORS

D. Glenar
Ionosphere Research Laboratory
The Pennsylvania State University
University Park, PA, 16802

T. Kostiuik, D.E. Jennings, and M.J. Mumma
Infrared and Radio Astronomy Branch, Code 693
NASA/Goddard Space Flight Center
Greenbelt, MD. 20771

ABSTRACT

A diode-laser based IR heterodyne spectrometer for laboratory and field use has been developed for high efficiency operation between 7.5 and 8.5 microns. The local oscillator is a PbSSe tuneable diode laser kept continuously at operating temperatures of 12-60K using a closed-cycle cooler. The laser output frequency is controlled and stabilized using a high-precision diode current supply, constant temperature controller, and a shock isolator mounted between the refrigerator cold tip and the diode mount. Single laser modes are selected by a grating placed in the local oscillator beam. The system employs reflecting optics throughout to minimize losses from internal reflection and absorption, and to eliminate chromatic effects. Spectral analysis of the diode laser output between 0 and 1 GHz reveals excess noise at many diode current settings, which limits the infrared spectral regions over which useful heterodyne operation can be achieved. System performance has been studied by making heterodyne measurements of etalon fringes and several Freon 13 (CF_3Cl) absorption lines against a laboratory blackbody source. Preliminary field tests have also been performed using the Sun as a source.

INTRODUCTION

To date, a number of studies have been made involving infrared heterodyne detection using tunable diode lasers (TDL's). The first effort was made in 1974 by Mumma et al. (ref. 1), where a limited tunability device and an 8-channel bank of 25 MHz filters were used to obtain blackbody continuum measurements of the Moon and Mars, and the detection of a laboratory N_2O line. Subsequent work included the detection of stratospheric ozone features by Frerking (ref. 2) and spectroscopy of C_2H_4 in the laboratory by Ku and Spears using a TDL heterodyne radiometer (ref. 3). Recently, Harward and Hoell have detected lines of several stratospheric species between 9 and 11 microns using a swept single channel TDL radiometer (ref. 4).

While there are a variety of problems associated with tunable diode lasers as local oscillators, they have the potential of continuous spectral tunability, which is not possible with gas laser systems. Current TDL technology allows coarse selection of operating wavelengths in regions between 4-20 μm , with single mode output powers marginally sufficient for useful heterodyne operation in spectral regions below 12 μm . The instrument discussed here has been developed as a general purpose TDL heterodyne spectrometer for high resolution ground based studies of the stratosphere and of astrophysical sources.

SPECTROMETER LAYOUT

A schematic diagram of the diode laser heterodyne spectrometer is shown in Figure 1. With the exception of the 50 percent reflecting - 50 percent transmitting ZnSe beamsplitter, which is used as a beam combiner, the use of off-axis parabolic mirrors provides an all reflecting system. This eliminates transmission losses and dispersion effects and improves the quality of focal images throughout the system.

The optical system is matched to the GSFC 48-inch telescope and an f/30 input beam. The incoming beam from the telescope results in a diffraction limited spot size (full width at half maximum) of about 300 microns at the telescope focus when the operating wavelength is near 8 microns. The telescope beam is collimated using a 30 inch focal length parabolic mirror and combined with the laser local oscillator beam at the beamsplitter. The combined beams are then demagnified and focussed onto the HgCdTe detector using f/4.5 collecting optics.

The TDL local oscillator is kept continuously at operating temperature (12-60K) using a closed cycle cooler and provides multi-mode output powers of a few mW near 1200 cm^{-1} , with a frequency separation of about 1.4 cm^{-1} between adjacent longitudinal modes. Using a 150 line/mm grating as a mode selector in the local oscillator beam results in a mode spacing of about 150 microns in the detector focal plane. This effectively isolates single modes of the laser L.O. The TDL is mounted so that the polarization of the E field vector is horizontal (in the plane of the figure). In this configuration the grating efficiency in first order exceeds 85 percent at wavelengths between 8 and 12 microns, and should be suitable regardless of the diode used.

A synchronous detection process is employed whereby a rotating blade chopper in combination with a stationary flat mirror provides signal and reference beams from the source in alternate half-cycles of the chopping cycle. By means of dual kinematic mirrors near the telescope focus, a blackbody reference source can be placed in either the signal or the reference beam for sensitivity measurements and instrument calibration. This blackbody source is also used for laboratory studies of subject gases by placing a gas cell in the blackbody beam.

In the remote observation mode, any remaining systematic errors, either in the telescope or in the synchronous detection scheme, can be removed by changing the signal and reference beam paths using a translatable dichroic mirror. The transmitted visible beam from this mirror also provides a conjugate image for guiding on remote sources.

HETERODYNE DETECTION ELECTRONICS

Two methods are now being employed for heterodyne detection using the instrument. Most of the initial heterodyne measurements were obtained in the manner shown in Figure 2. Here the total I.F. signal from the IR detector is amplified (50 dB) using a low-noise wideband (0-1.5 GHz) amplifier. A portion of this output is again amplified and band limited using a 150 MHz low pass filter. The filter output is square-law detected and the D.C. signal is delivered to a lock-in amplifier and recorder. Using a high precision TDL current supply, the diode laser can be slowly tuned to scan the desired spectral region. The current supply delivers a voltage to the recorder X-axis which is proportional to the diode current. This produces a single channel scan of heterodyne signal as a function of diode current. By repeating the scan using a calibration etalon, a relative frequency scale can be assigned to the X-axis. The I.F. bandwidth, 150 MHz in this configuration, is the filter cutoff frequency. The wavelength resolution is therefore roughly 300 MHz and the effective integration time is the time constant setting on the lock-in amplifier, typically 1.25 seconds.

Recently, an RF filter bank spectral line receiver has been added to the system which provides multiplexing and time integration capability for measurements on weak astronomical sources. Using this method, the I.F. output is split into 32 contiguous frequency channels each 25 MHz wide, for a total coverage of 800 MHz in the I.F. This block of filters can be steered to any portion of the I.F. by mixing with a second RF local oscillator. The output of each filter is then digitized synchronously with the chopper and the output can be displayed and printed after each integration period. A variation of this spectral line receiver design is discussed in detail by Buhl and Mumma (ref. 5) and Mumma et al. (ref. 6). In the future this spectral line receiver will be extended to 1600 MHz coverage at 25 MHz resolution and a separately tunable high resolution filter bank of 32, 5 - MHz filters will be added.

DIODE LASER LOCAL OSCILLATOR

The TDL oscillator currently installed in the instrument is a PbSSe device manufactured by Laser Analytics, Inc. The output power in a single mode from this laser has been measured to be as large as 1 mW, although typical single mode powers for heterodyne work lie below 200 μ W. The operating frequency of this device varies from about 1170 cm^{-1} at a diode current of 0.20 A and mount temperature of 12 K, to nearly 1290 cm^{-1} at 2.0 A and 60 K.

The TDL is mounted on a vibrationally and thermally isolated platform that is enclosed in an evacuated shroud, as illustrated in Figure 3. This design scheme is discussed by Jennings and Hillman (ref. 7) and has been used successfully to stabilize a diode for direct absorption spectroscopic studies. The laser and copper mount are maintained at cryogenic temperatures by coupling the diode mount through a flexible braid to the second stage of a two-stage closed-cycle helium refrigerator. This braid, together with a metal bellows on the vacuum shroud isolate the diode mount assembly from mechanical shock associated with cooler operation. The

entire assembly is enclosed in a 50 K radiation shield attached to the refrigerator first stage. This serves to lower the minimum mount temperature by reducing the conducted heat load on the second stage. A high-precision temperature controller is used to preset the diode mount temperature anywhere between the cooling limit of 11 K, and the maximum recommended diode temperature of about 70 K. With a properly positioned temperature feedback loop, consisting of a sensor and heater at the diode mount, the temperature has been stabilized to within ± 1.0 mK over a 1 hour period. Since the temperature tuning rate for our present diode has been found to be approximately 10 MHz/mK, this corresponds to a frequency stability over the same time period of about ± 10 MHz. Stability better than ± 5 MHz is expected over shorter time periods.

Once the temperature has been set, the diode output frequency is fine tuned by varying the diode current, as previously described, or the diode frequency may be preset to some fixed value. The PbSSe device now in use exhibits a current tuning rate of about $1.2 \text{ cm}^{-1}/\text{A}$ near 1200 cm^{-1} . This rate is slow enough so that contributions to the frequency drift from current supply instabilities are negligible.

INITIAL RESULTS

Figure 4 illustrates the current tuning procedure applied to a single channel scan of a 15 percent absorbing feature, near 1180 cm^{-1} . The results of two heterodyne scans are shown, obtained by first placing a 10 cm cell containing 25 Torr of Freon 13 in front of the blackbody and then replacing the cell with a 1 inch etalon. Roughly $150 \mu\text{W}$ of single mode laser power was available when these measurements were made. The inset at the upper right shows direct absorption measurements that were made by placing the same two objects in the local oscillator beam and current tuning again over the same region. The baseline in both heterodyne and direct scans refers to the absorption feature, and not to the etalon fringe pattern which has been shifted down in both cases and is used here to establish a relative frequency scale. The lack of any observable structure in the absorption line during the heterodyne scan of the type seen in the direct mode is due mainly to the system resolution of 300 MHz, comparable with the line spacing in the multiplet.

The best achievable double side-band signal to noise ratio (SNR) in this experiment was found to be about 250, for an IF bandwidth of 150 MHz, integration time of 1.25 sec, and blackbody temperature of 1273 K. The expected SNR for an ideal system under these same conditions is about 19,000, making the measured system Δ (ref. 6) about 80 in this region of the scan. A more typical value of achievable SNR is 140 in the noisier portion of the scan, which yields a Δ of approximately 145. This increased noise, which turns on at the same current value in both heterodyne scans has been found to originate in the diode laser output. Studies using a spectrum analyzer show that this noise results from the generation of small satellite modes close enough in frequency to the primary mode to be seen in the I.F. These small modes show amplitude modulations that are correlated with vibrations of the diode cooler second stage, which is mounted on the optical bench. As might be expected, the

noise superimposed on the resulting spectrum is strongly time coherent with a frequency of about 2.5 Hz, which is the cycling period of the refrigerator second stage. Many of the modes which are otherwise suitable for heterodyne detection are much more sensitive to cooler vibrations than the example shown here, so that a large fraction of the available diode tuning range is made unsuitable for heterodyne work with the present system. Plans are being made to install a new diode in the near future. Test results obtained at that time will reveal whether the diode we are currently using is unusually vibration sensitive, or whether further de-coupling of the diode from the closed cycle cooler is required.

After interfacing the spectrometer with the R.F. spectral line receiver, the same gas cell with 40 Torr of Freon 13 was placed in the beam of a 1270 K blackbody, producing the results shown in Figure 5. 25 RF channels of 25 MHz width were used in these preliminary measurements. The features shown here lie in the same wavelength region (near 1180 cm^{-1}) as the weaker lines discussed earlier. Both spectra in Figure 5 were obtained using the same diode laser mode and a total power of about $200\mu\text{W}$ on the detector. By first locating the lines in direct absorption and tuning the diode to the line center position, the line could be positioned properly for heterodyne detection by shifting the diode current a predetermined amount. In so doing, the feature is moved an appropriate distance out in either sideband from the zero of I.F. Each result consists of an average of four consecutive runs for a total effective integration time of 3.2 minutes. The spectrum is divided by the average of four scans taken during the same time period with the gas cell removed, yielding a quotient spectrum with fractional transmittance as the ordinate. This eliminates the effect of gain variations with frequency from the amplifiers and detector.

The instrument has very recently been placed in operation at the Coude focus of the 48 inch reflecting telescope at the Goddard optical site and additional 25 MHz filters were added to the filter bank for a total I.F. coverage of 800 MHz.

The first observational test performed was a comparative measurement of detector response in the heterodyne mode using the Sun and a local blackbody source at a number of temperature settings. Figure 6 is a composite of these blackbody measurements, taken sequentially, with all system parameters held constant during the duration of the tests. Each measurement was made with an integration time of about 45 seconds and an LO power of $\sim 200\mu\text{W}$. The modulations observed in each scan are caused by standing waves in the RF cables and can be removed by ratioing as was done in Figure 5.

The positive slope for all curves is predominantly due to the detector and preamplifier roll-off at high I.F. The ratios of heterodyne signals for any two temperatures have been found to be close to the expected values at the high frequency end. Closer to the zero of I.F., however, the high temperature curves grow anomalously large. This effect is due to the absence of an infrared filter in the blackbody (or Sun) beam. For the temperatures of interest here, a large fraction of the total energy flux is converted to shot noise power by the detector and synchronously detected. Since the bolometric (T^4) term increases much more quickly with temperature than the heterodyne signal term at these wavelengths, the synchronously detected blackbody shot noise becomes large compared to the signal as the

source temperature increases. This result emphasizes the importance of a relatively narrow band filter in the signal beam when making measurements against hot sources.

SUMMARY

These results show that diode laser heterodyne spectroscopy has evolved to the point where remote sources other than the Sun can be detected with adequate signal-to-noise. From the level of system performance achieved so far at 6.8 microns, one can expect signal to noise ratios in excess of 10 at $> 10^6$ resolving power for a continuum temperature of 250 K and less than 30 minutes integration time, making planetary spectroscopic observations feasible. In addition, grating and Fabry-Perot spectrometer studies of several compact HII regions reveal large fluxes in the fine structure lines of SIV and ArIII near 10.51 and 8.99 microns respectively. Calculations show that these features, if narrow in linewidth, should be easily observable for integration times of less than 1/2 hour.

One of the major remaining problems is the lack of a satisfactory method of absolute frequency calibration to within the accuracy required. Careful calibration of an IR monochromator allows a frequency determination to within about $\pm 0.1 \text{ cm}^{-1}$. This uncertainty is equal to about $\pm 3 \text{ GHz}$, which is still considerably broader than the 800 MHz filter bank recently interfaced with the system, and more than two orders of magnitude broader than a single resolution element of 25 MHz.

A supplemental scheme for frequency calibration uses a transfer standard (gas) having numerous lines of established rest frequencies in the region of interest. By interpolating between the diode current settings at the centers of these standard lines, the operating wavelength may be established with high precision. In the 1200 cm^{-1} region, the recently available N_2O calibration standards can produce the desired $< 0.001 \text{ cm}^{-1}$ calibration accuracy. (Data from work (as yet unpublished) by W. B. Olson, A. G. Maki, and W. J. Lafferty, National Bureau of Standards, Gaithersburg, Md.)

References

1. Mumma, M., T. Kostiuik, S. Cohen, D. Buhl and P.C. Von Thuna, Heterodyne spectroscopy of astronomical and laboratory sources at $8.5\mu\text{m}$ using diode laser local oscillators, Space Science Reviews, 17, 661-667, 1975.
2. Frerking, M.A. and D.J. Muehlner, Infrared heterodyne spectroscopy of atmospheric ozone, Applied Optics, 16(3), 526-528, 1977.
3. Ku, R.T. and D.L. Spears, High-sensitivity heterodyne radiometer using a tunable diode laser local oscillator, Solid State Research, 1, 1, 1977.
4. Harward, C.N. and J.M. Hoell, Atmospheric solar absorption measurements in the 9-11 micron region using a diode laser heterodyne spectrometer. Heterodyne Systems and Technology, NASA CP-2138, 1980. (Paper 16 of this compilation.)
5. Buhl, D. and M. J. Mumma, RF spectrometers for heterodyne receivers. Heterodyne Systems and Technology, NASA CP-2138, 1980. (Paper 28 of this compilation.)
6. Mumma, M.J., T. Kostiuik, D. Buhl, A $10\mu\text{m}$ laser heterodyne spectrometer for remote detection of trace gases, Optical Engineering, 17, 50, 1978.
7. Jennings, D.E. and J.J. Hillman, Shock isolator for diode laser operation on a closed-cycle refrigerator, Review of Scientific Instruments, 48(12), 1568-1569, 1977.

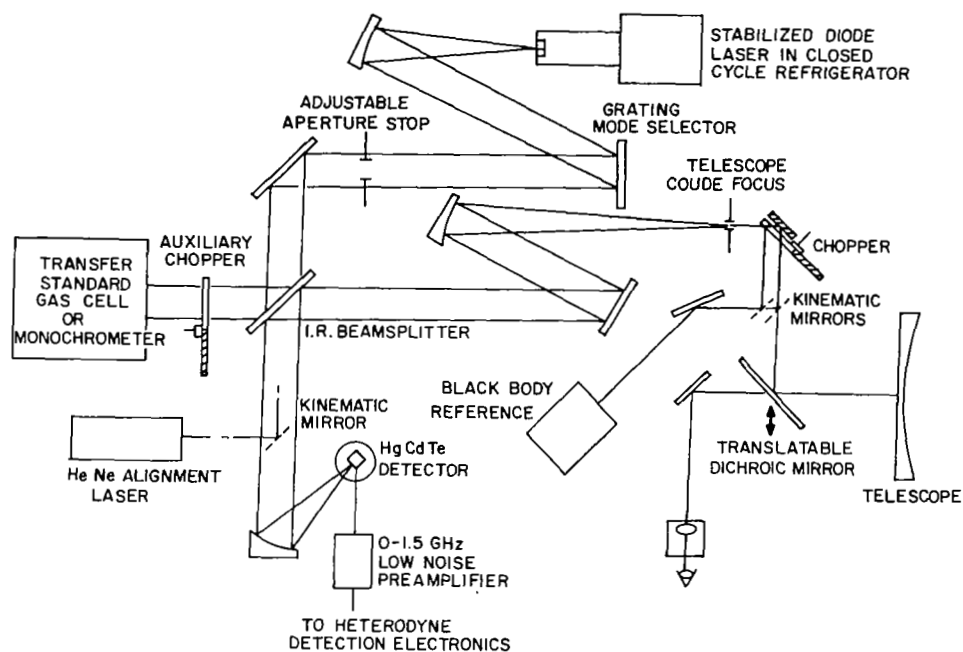


Figure 1.- Schematic diagram of diode laser heterodyne spectrometer optical front end.

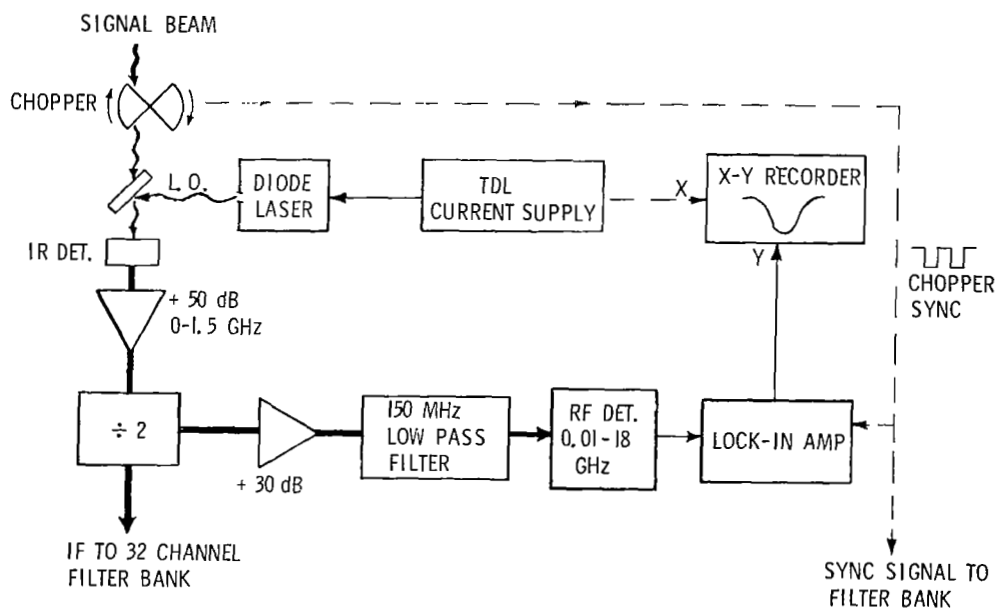


Figure 2.- Diagram of current tuning method for producing heterodyne spectra.

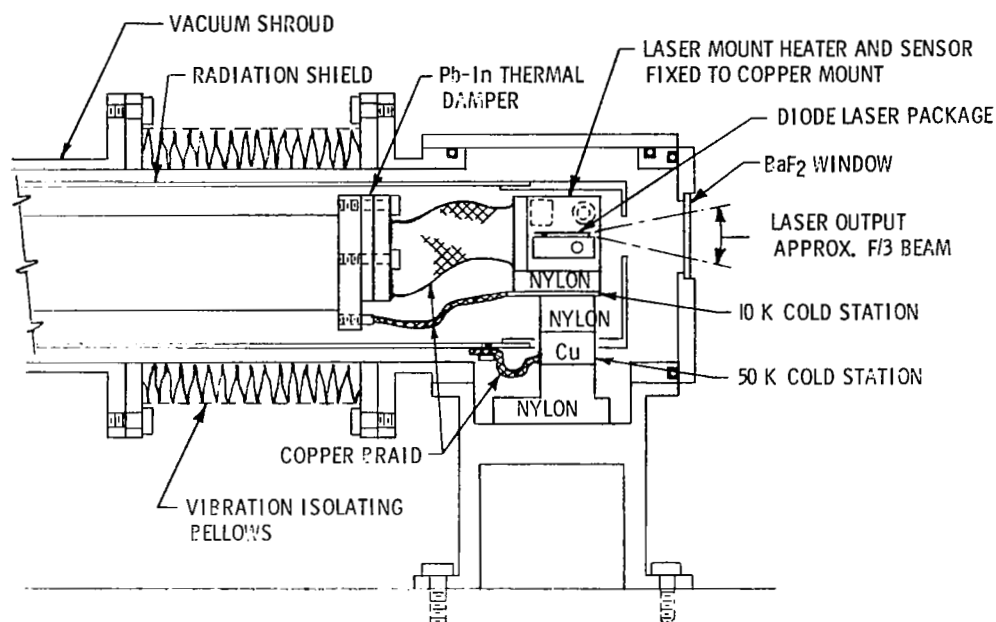


Figure 3.- Diode laser mounting scheme.

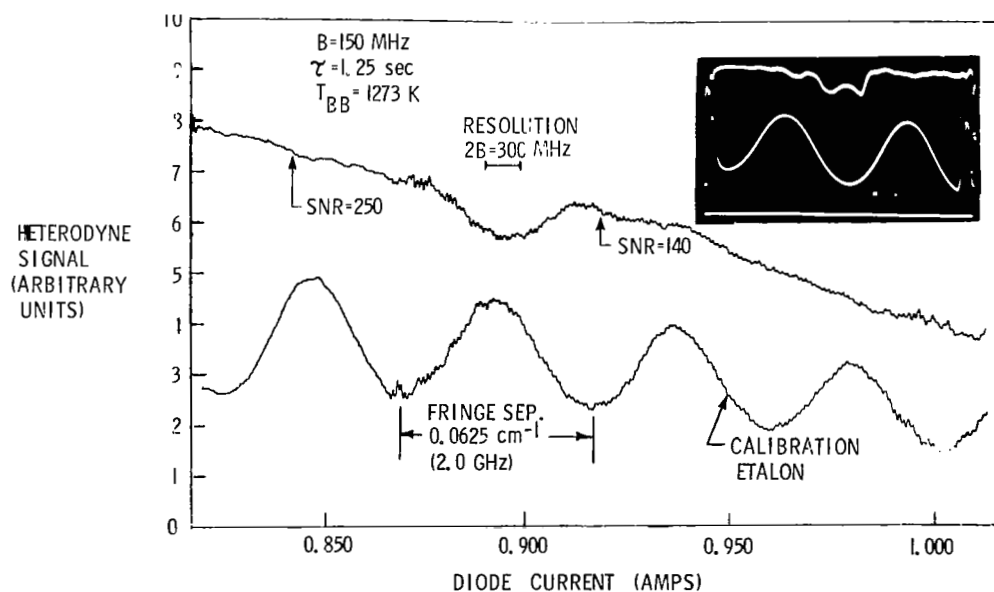


Figure 4.- Single channel heterodyne scan and direct absorption scan (inset) of Freon 13 multiplet and 1 inch etalon near 1180 cm^{-1} . The baseline of the etalon scan has been shifted for clarity. The etalon scans are shown below the CF_3Cl absorption scans in each case.

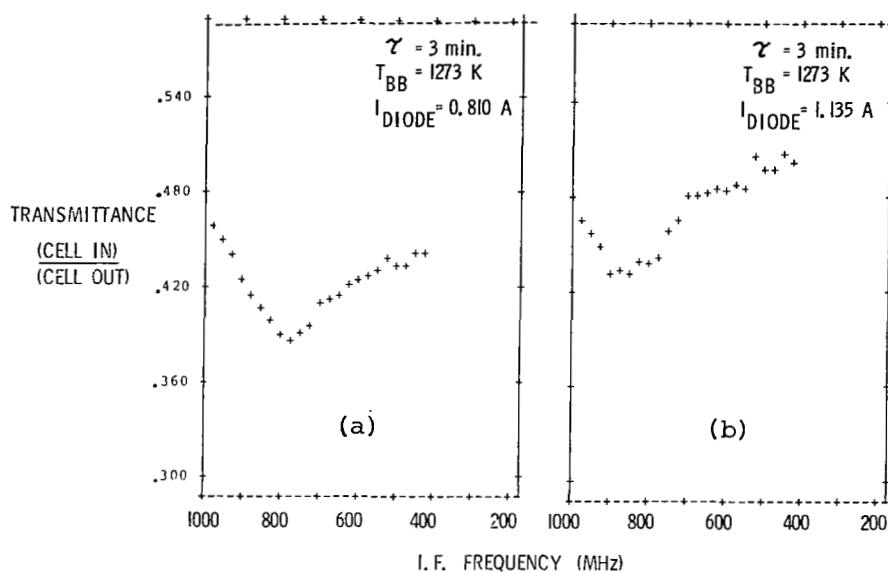


Figure 5.- Heterodyne measurements of pressure broadened Freon 13 (CF_3Cl) lines near 1180 cm^{-1} using a 23 channel RF spectral line receiver and 10 cm absorption cell. Trace (a) shows a blend of several discrete lines while trace (b) is an isolated feature.

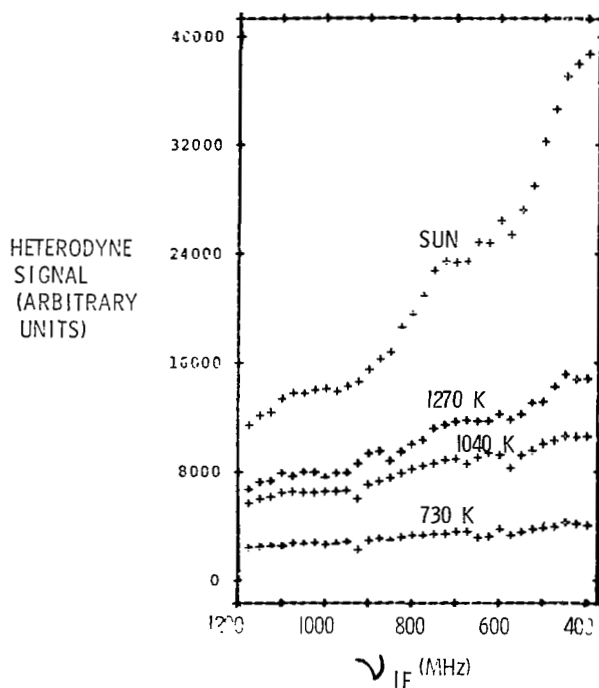


Figure 6.- Preliminary blackbody heterodyne efficiency measurements using spectral line receiver.

ATMOSPHERIC SOLAR ABSORPTION MEASUREMENTS IN
THE 9-11 MICRON REGION USING A DIODE
LASER HETERODYNE SPECTROMETER

Charles N. Harward
Department of Physics
Old Dominion University

Norfolk, VA 23508

and

James M. Hoell, Jr.

Langley Research Center

Hampton, VA 23665

ABSTRACT

A tunable diode laser heterodyne radiometer has been developed for ground-based measurements of atmospheric solar absorption spectra in the 9 to 12 micron spectral range. The performance and operating characteristics of this Tunable Infrared Heterodyne Radiometer (TIHR) will be discussed along with recently measured heterodyne solar absorption spectra in the 10-11 micron spectral region.

INTRODUCTION

Fixed frequency or discretely tunable CO₂ laser local oscillators (LO) have been used in heterodyne radiometers to measure selected species of atmospheric or astronomical interest (1), (2), (3), (4), (5), (6). The range of applications have in general depended upon coincidences between CO₂ laser emission wavelengths and absorption lines of the desired species. Often no overlap exists or the close coincidences that are available are contaminated by interferences from other atmospheric species. Far more important, however, for monitoring a wide range of species, is the limited overall coverage available from conventional CO₂ laser systems (i.e., from ~ 9 to 11.5 micrometers). Clearly, the use of passive IR heterodyne techniques would be greatly enhanced through the use of tunable laser LO's operating over a wider spectral region.

Tunable Diode Lasers (TDL) are currently the only commercially available devices that offer the potential for spectral coverage from 3-30 micrometers. These lasers have been successfully used as LO's in heterodyne systems for astronomical observations (7) of the thermal emission from Mars and the moon and for limited stratospheric measurements of solar absorption by ozone (8). Ku and Spears (9) have reported on the use of TDL LO to detect C_2H_4 in a laboratory environment using a 900K blackbody as the source and they also report on excess RF noise generated in a heterodyne system by a diode laser. Additional laboratory studies on the effects of TDL generated noise on heterodyne applications have been reported by (10), (11), (12), and (13). Of particular significance are the results reported by Allario et al (12) demonstrating that the performance of a TDL radiometer can be comparable to that available from a CO_2 laser system.

In this paper, we report on the design and use of a TDL heterodyne radiometer for high resolution ground based atmospheric solar absorption measurements. The performance and operating characteristics of the system will be discussed along with atmospheric absorption of HNO_3 , O_3 , CO_2 , and H_2O in the 9-11 micrometer spectral region. These data along with additional spectra to be obtained with this system, when coupled to existing Fourier transform spectra (14) will be useful for identifying optimum LO wavelengths for future tropospheric and stratospheric monitoring.

EXPERIMENTAL DETAILS

The Tunable Infrared Heterodyne Radiometer (TIHR) discussed here is shown in figure 1. It is currently coupled to an eight inch heliostat located on the roof of a laboratory at Langley Research Center, Hampton, VA. The current version of the TIHR is designed to operate over the spectral region from 8 to 12 microns. The lower wavelength cutoff is dictated by a long pass optical filter in the solar path, while the long wavelength cutoff is due to the spectral response of the high speed HgCdTe photomixer. The TDL's are mounted in a commercially available closed cycle cooler modified to reduce frequency fluctuations in the laser output. Radiation from the TDL is collected and collimated by an f/1 Ge lens. The monochromator shown in figure 1 is only used for coarse wavelength characterization and identification. During operation of the radiometer, the monochromator is bypassed by using the first flip mirror and for multimode laser operation, a scanning etalon is inserted in the optical path to isolate and track a single TDL mode. The TDL and solar radiation are combined at the 1:1 ZnSe beamsplitter and then focused onto a LN_2 cooled HgCdTe photomixer with an f/2 Ge lens. The incoming solar radiation from the heliostat (or 1300K blackbody for calibration and alignment) is mechanically chopped and filtered using the eight micron long pass optical filter. The TDL radiation reflected by the beamsplitter is directed into a wavelength identification system similar to that used in standard TDL spectroscopic systems. The gas cell, with a suitable reference gas, provides an absolute frequency scale while the solid Ge etalon provides a relative frequency scale to account for the nonlinear tuning rate of the laser over

extended wavelength regions and between absorption lines from the reference gas.

One of the advantages of tunable LO's is that the radiometer can be used either as a scanning or fixed frequency system. As a fixed frequency system, the TIHR utilizes a wide bandwidth photomixer and IF amplifier (i.e., 5-2000 MHz). The IF output is channelized using up to 16 IF filters. The output from each filter is rectified by a crystal detector and synchronously detected by a lock-in amplifier referenced to the chopper in the solar path. In the scanning mode reported here, the wideband width amplifier is replaced with a lower noise amplifier, having a much narrower bandwidth (i.e., 1-200 MHz), a single 100 MHz bandwidth low pass IF filter, crystal detector and lock-in amplifier. The scanning mode offers the advantage of wider spectral coverage limited only by the tuning range of the TDL. The fixed frequency operation, however, offers a multiplex advantage through the use of multiple IF channels which can be important for quantitative measurements of trace molecules.

It is well known that the gain curve of lead salt lasers is broad enough to sustain several longitudinal modes in its emission spectrum. For most applications, a single longitudinal mode is desired. To facilitate the use of the scanning mode of the TIHR, the etalon shown in figure 1 was used to isolate and track a single longitudinal mode. Use of an etalon for mode isolation provides a simple compact technique with high throughput. Furthermore, the center of the bandpass curve for the etalon can easily be locked to the peak of an individual mode for operation over an extended spectral range. The electronic feedback system used for tracking the TDL was assembled using standard components. After centering the etalon transmission peak on the desired TDL mode, a small 1000 Hz AC dither voltage obtained from the internal oscillator of a lock-in amplifier was applied to the etalon control voltage. Near the etalon transmission peak the TDL radiation is modulated such that the modulation amplitude is proportional to the offset between the peak of the etalon transmission curve and the isolated TDL mode; the phase of the modulated TDL signal relative to the dither voltage depends on which side of the etalon transmission curve the isolated mode is located. Therefore, a phase/amplitude dependent signal is obtained by synchronously detecting the modulation on the DC current generated in the photomixer by the TDL radiation. This is used to control the etalon plate spacing via a variable gain high voltage amplifier internal to the etalon control module. To minimize the effects on the heterodyne signal, the dither frequency was chosen to be an odd multiple of the chopper frequency in the solar path and much larger than the effective post integration frequency. Figures 2a and 2b illustrate the operation of the tracking etalon. The upper trace in each figure is the heterodyne signal obtained using a 1300K blackbody source while the lower trace is the output of the solid etalon in the wavelength identification section. In figure 2a, the general shape of the heterodyne signal results from the convolution of the TDL power curve and the bandpass function of the static etalon. In figure 2b, where the etalon is locked to the peak of the TDL mode, there is little variation in the TDL power curve over the same tuning range used in figure 2a. More important, the continuous spectral coverage is now limited only by the tuning range of the individual mode and the overall signal variation is now due to the true variation in laser power.

ATMOSPHERIC SOLAR ABSORPTION

Table I gives the spectral regions that have been covered to date using two lead salt lasers; one provided by Dr. Wayne Lo, from General Motors Research Laboratory, and the other obtained from Laser Analytics, Inc. Figures 3, 4, and 5 show spectra in which all the features have been identified. For the data shown here, the LO power ranged from about 200 to 300 microwatts. Each spectra was recorded in the scanning mode with a 100 MHz low pass IF filter (i.e., spectral resolution $\sim .007 \text{ cm}^{-1}$) and a post integration time of 2.5 secs. The measured signal-to-noise ratio (SNR) ranged from 200 to 300. For the above system parameters, the maximum expected SNR viewing the unattenuated solar energy is approximately 800. Atmospheric attenuation, mismatch between the TDL and signal wavefronts along with uncertainties in some system parameters such as optical throughput and photomixer efficiency explain the lower SNR actually measured. The data shown in figure 3 includes the atmospheric solar absorption spectra from 921.1 cm^{-1} to 921.6 cm^{-1} ; the absolute frequency mark was obtained from the wavelength ID gas cell with ammonia as the calibration gas, and the relative frequency marks from the solid germanium etalon. The only feature seen in the atmospheric spectra is attributed to water vapor. It is interesting to note that for this spectral region atmospheric ammonia should be discernible at tropospheric levels greater than 1 ppb. Also, note that the halfwidth of the water vapor absorption is approximately 3 GHz as expected for a tropospheric species. Figure 4 shows a spectral region exhibiting ozone absorption features with halfwidths which can be attributed to stratospheric pressure broadening and a CO_2 feature with a halfwidth consistent with tropospheric broadening. The absolute wavelength for this spectra was obtained from the R(34) ($0001 - 0200$) CO_2 feature. This spectrum also shows the affect of a longitudinal mode change in the TDL emission which occurs at the point separation regions 1 and 2 where region 1 is near 1086.8 cm^{-1} and region 2 is near 1088.6 cm^{-1} . Figure 5 shows the atmospheric absorption spectra from about 896.0 cm^{-1} to 896.6 cm^{-1} , along with two reference spectra from the wavelength identification section of the TIHR. The reference spectra is for pure HNO_3 (upper trace) and low pressure HNO_3 broadened by approximately 50 torr of air (middle trace). The fine structure seen in the atmospheric spectra can be attributed to stratospheric nitric acid. The wavelengths identified on this figure were obtained from Brockman et al. (15). The broad feature at 896.507 cm^{-1} has been attributed to absorption primarily by H_2O and CO_2 . It is significant to note that lower resolution ($.02 \text{ cm}^{-1}$) ground based data (14) available for this region shows only the H_2O and CO_2 absorption and little if any of the HNO_3 fine structure seen in figure 5. Figure 6 shows a synthetic atmospheric spectra generated using the line parameters for CO_2 , H_2O , (16) and HNO_3 (15). With the exception of the overall shape of the experimental data due to the TDL power variation, the similarity between the synthetic spectrum and the experimental spectrum is evident. With the addition of a blackbody calibration source, the experimental data could be used to obtain HNO_3 profiles.

SUMMARY

In summary, we have shown the design and operating characteristics of a tunable infrared heterodyne radiometer which uses a tunable diode laser as a local oscillator. The use of a scanning etalon as a mode isolator as well as the improvement in system performance when the etalon was locked to the TDL was shown. Atmospheric absorption spectra of selected regions from the 9 and 11 micrometer were presented. These data exhibited the highest resolution HNO_3 atmospheric spectra obtained to date and demonstrated the ability of the TIHR to discriminate between absorption which occurs in different atmospheric regions.

REFERENCES

1. Abbas, Mian M.; Kosluk, Theodore; Mumma, Michael J.; Buhl, David; Kunde, Virgil G.; and Brown, Larry W.: Stratospheric Ozone Measurement with an Infrared Heterodyne Spectrometer. *Geophy. Res. Lett.*, 5, 317-320, 1978.
2. Betz, A. L.; Johnson, M. A.; McLaren, R. A.; and Sutton, E. C.: Heterodyne Detection of CO₂ Emission Lines and Wind Velocities in the Atmosphere of Venus, *Astrophys. J.*, 208, L141, 1976.
3. Hoell, J.; Harward, C.; and Williams, B.: Remote Infrared Heterodyne Radiometer Measurements of Atmospheric Ammonia Profiles. *Geophy. Res. Lett.*, 7, 1980.
4. Menzies, Robert T.: Atmospheric Monitoring Using Heterodyne Detection Techniques. *Optical Eng.*, 17, 44-49, Jan. - Feb. 1978.
5. Mumma, Michael J.; Kostouk, Theodor; Buhl, David: A 10 μ m Laser Heterodyne Spectrometer for Remote Detection of Trace Gases. *Optical Eng.*, 17, 50-55, Jan. - Feb. 1978.
6. Peyton, B. J.; Lange, R. A.; Savage, M. G.; Seals, R. K.; and Allario, F.: AIAA Paper 77-73, AIAA 15th Aerospace Sciences Meeting, Los Angeles, CA, Jan. 24-26, 1977.
7. Mumma, M.; Kostiuik, I.; Cohen, S.; Buhl, D.; and von Thuna, P. C.: Infrared Heterodyne Spectroscopy of Astronomical and Laboratory Sources at 8.5 μ m. *Nature*, 253, 514-516, Feb. 1975.
8. Frerking, M. A.; and Muehlner, D. J.: Infrared Heterodyne Spectroscopy of Atmospheric Ozone. *Appl. Opt.*, 16, 526-528, 1977.
9. Ku, R. T.; and Spears, D. L.: High Sensitivity Infrared Heterodyne Radiometer Using a Tunable-Diode Laser Local Oscillator, *Optics Lett.*, 1, 84-86, 1977.
10. Harward, C.: The Evaluation of a HgCdTe Photomixer with a Tunable Diode Laser (TDL) and the Evaluation of TDL's as a Local Oscillator in a Heterodyne Detection System. Final Report for NASA Grant NSG 1197, Technical Report PGSTR-AP 77-63, Nov. 1977.
11. Harward, C.; and Hoell, J.: Optical Feedback Effects on the Performance of Pb_{1-x}Sn_xSe Semiconductor Lasers. *Appl. Opt.* 18, 3978, 1979.
12. Allario, F.; Hoell, Jr.; Katzberg, S.; and Larsen, J.: An Experimental Concept to Measure Stratospheric Trace Constituents by Laser Heterodyne Spectroscopy. Accepted for publication in *Applied Physics*, 1980.

13. Savage, M.; Augeri, R.; and Peyton, B. J.: Application of Tunable Diode Lasers to Infrared Heterodyne Spectroscopy. Second Conference on Lasers and Applications, Orlando, FL, Dec. 17-21, 1979.
14. Goldman, A.; and Blatherwick, R. D.: New Atlas of IR Solar Spectra. National Science Foundation, Annual Report for Grant ATM 76-83908, 1978.
15. Brockman, P.; Bair, C.; and Allario, F.: High Resolution Spectral Measurement of the HNO_3 11.3 μm Band Using Tunable Diode Lasers. Appl. Opt., 17, 91, 1978.
16. McClutchey, R.; Benedict, W.; Clough, S.; Burch, D.; Calfee, R.; Fox, K.; Rothman, L.; and Garing, J.: AFCRL Atmospheric Absorption Line Parameter Compilation. Environmental Research Paper 434, AFCRL-TR-73-0096, Jan. 1973.

Table 1. Approximate wave number range and atmospheric species which cause the absorption.

<u>WAVE NUMBER RANGE (cm⁻¹)</u>	<u>ABSORBING SPECIES</u>
1064.0 - 1064.6	O ₃
1066.2 - 1067.1	O ₃ , CO ₂ , H ₂ O
1068.3 - 1069.4	O ₃ , CO ₂
1070.9 - 1071.7	O ₃
1072.8 - 1073.7	CO ₂ , O ₃
1075.0 - 1076.0	O ₃ , H ₂ O
1077.4 - 1078.2	O ₃ , CO ₂
1079.4 - 1080.2	O ₃ , CO ₂
1081.8 - 1082.5	CO ₂ , O ₃
1084.2 - 1085.0	CO ₂ , O ₃
1086.5 - 1087.2	CO ₂ , O
1088.7 - 1089.3	O ₃ , CO ₂
1091.2 - 1091.8	H ₂ O, O ₃
1092.7 - 1093.2	O ₃
896.0 - 896.6	HNO ₃ , CO ₂ , H ₂ O
921.1 - 921.6	H ₂ O, NH ₃

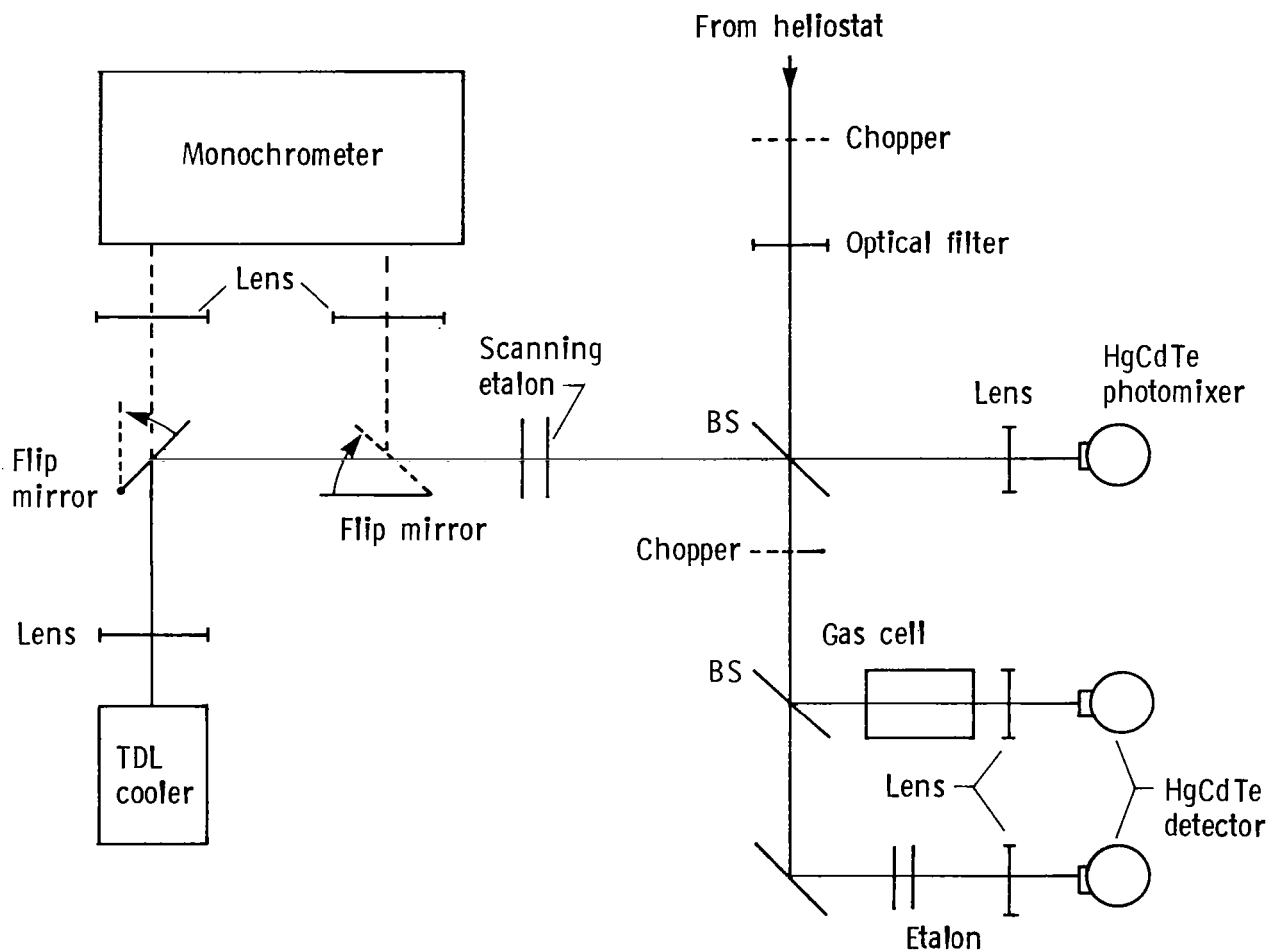
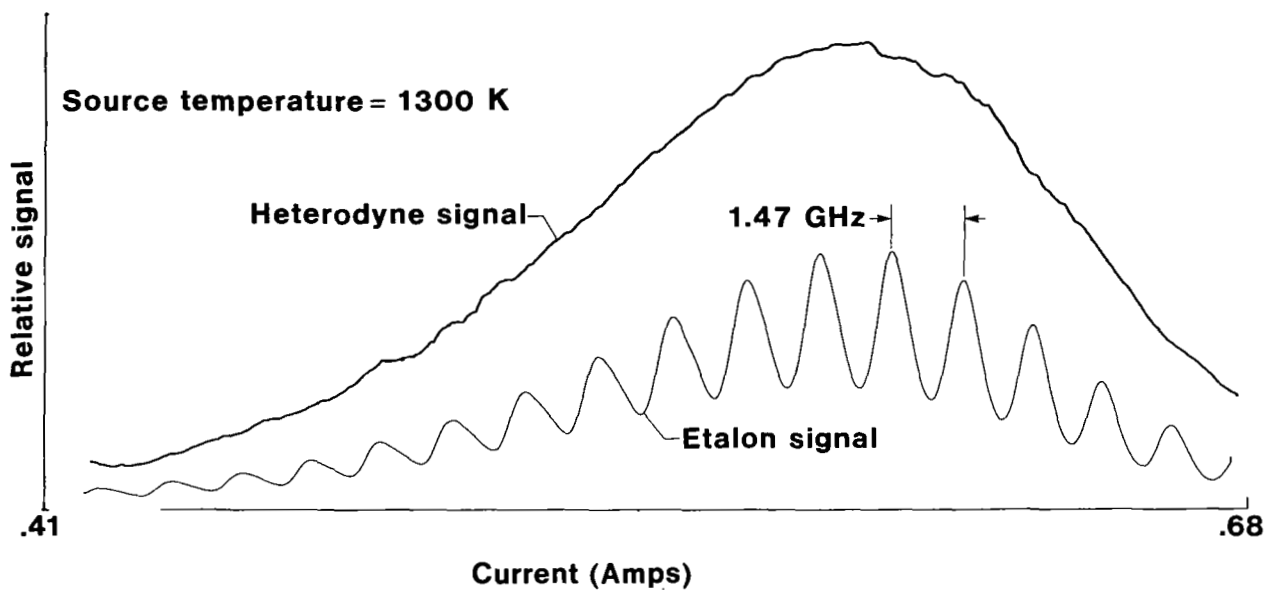
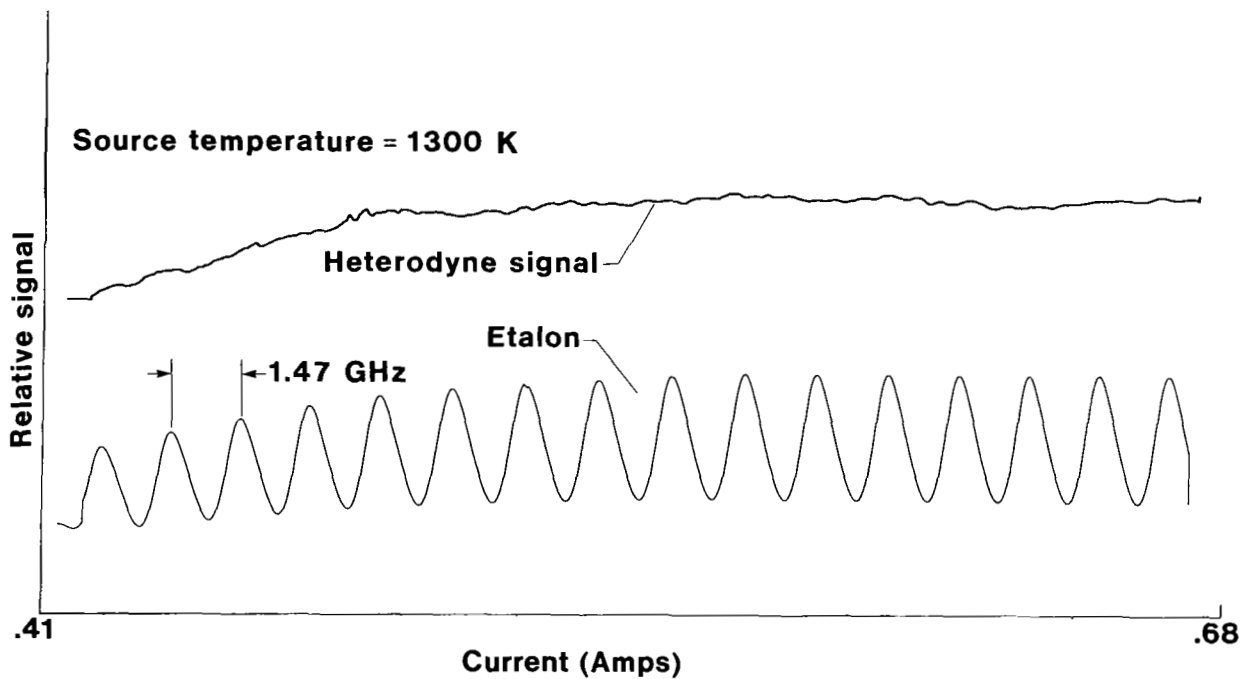


Figure 1.- Optical schematic of the tunable infrared heterodyne radiometer.



(a) Blackbody (1300K) heterodyne signal obtained with a TDL mode isolated with a fixed spacing on the scanning etalon in figure 1.



(b) Same as in (a) above except the scanning etalon was locked to the TDL mode using a conventional feedback circuit.

Figure 2.- Effect of TDL mode isolation technique on blackbody heterodyne signal.

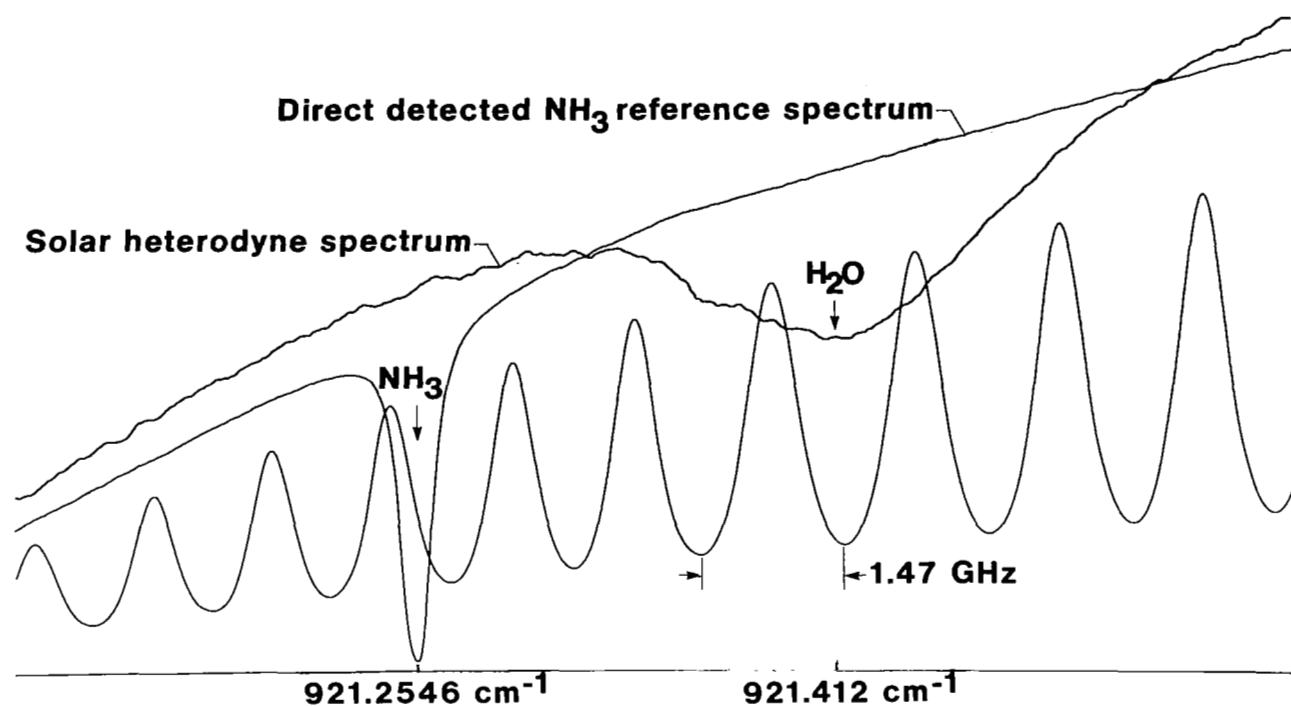


Figure 3.- Heterodyne solar absorption spectrum of atmospheric water vapor with NH_3 as the reference gas.

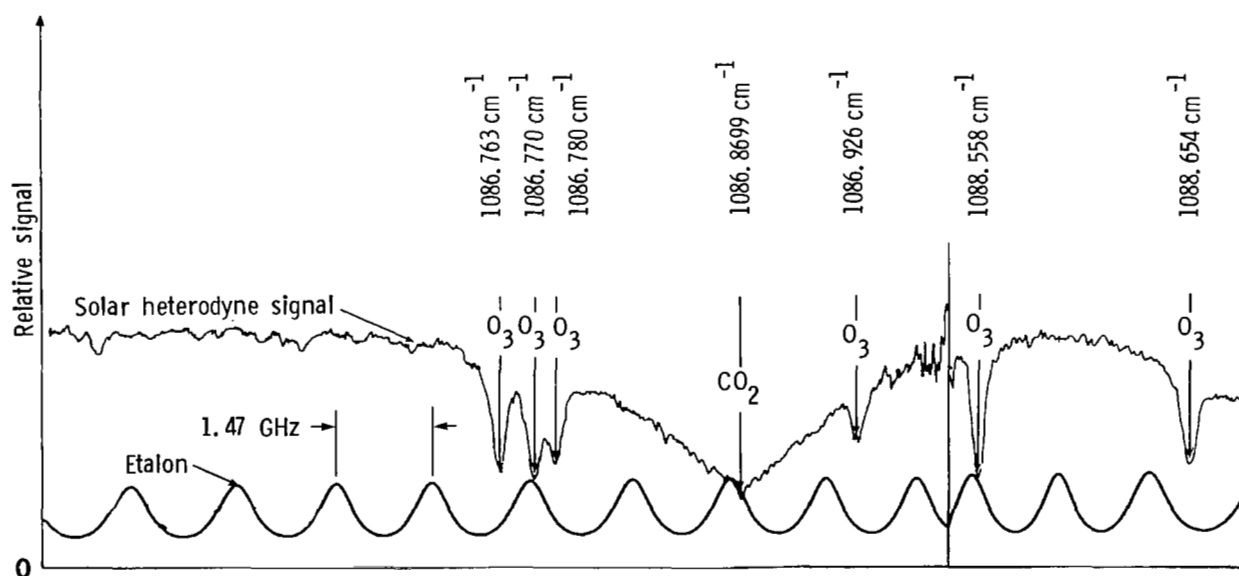


Figure 4.- Heterodyne solar absorption spectra of atmospheric ozone and CO_2 . The absolute wavelength was identified by the broad atmosphere $\text{R}(34) (00^0_1 - 02^0_0) \text{CO}_2$ line.

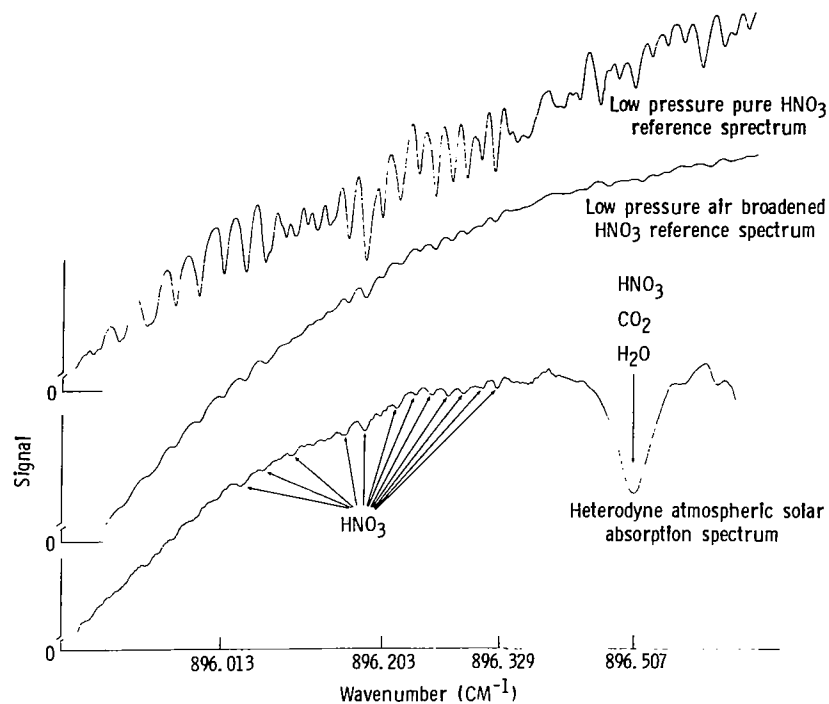


Figure 5.- Heterodyne solar absorption spectrum compared to two HNO₃ reference spectra.

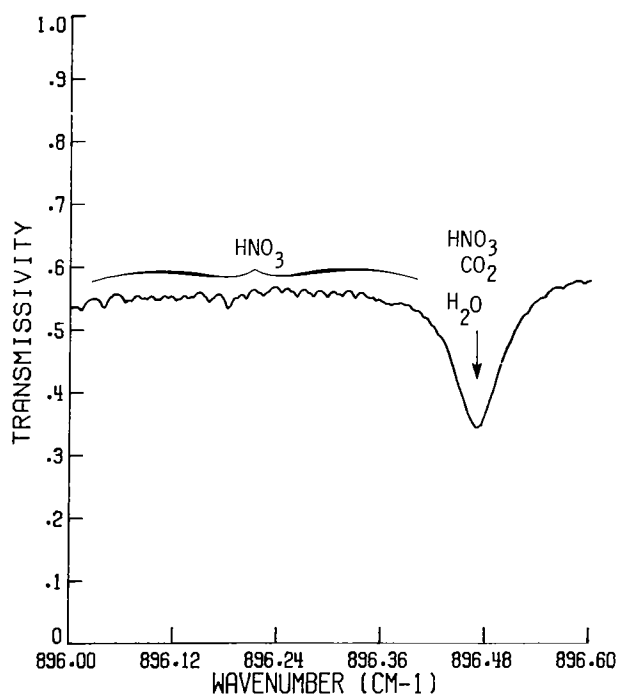


Figure 6.- Synthetic atmospheric spectrum in the same spectral region as shown in figure 5.

SENSITIVITY STUDIES AND LABORATORY MEASUREMENTS FOR
THE LASER HETERODYNE SPECTROMETER EXPERIMENT

Frank Allario and Stephen J. Katzberg
Langley Research Center

Jack C. Larsen
Systems and Applied Sciences, Corp.

INTRODUCTION

Several experiments involving spectral scanning interferometers (ref. 1) and gas filter correlation radiometers (ref. 2) using limb scanning solar occultation techniques are currently under development for measurements of stratospheric trace gases from Spacelab and satellite platforms. The scientific requirements for improving the current understanding of the physics and chemistry of the upper atmosphere were recently summarized by several NASA-sponsored science working groups (refs. 3, 4), in which the importance of direct measurements of radical species in the stratosphere were highlighted for the three basic chemical families including chlorine, hydrogen, and nitrogen. Recently, NASA sponsored a science working group to study the role of high resolution spectroscopic measurement techniques in the future development of the Upper Atmosphere Research Program (ref. 5), and emphasized the importance of future collaborative experiments between spectrally scanning interferometers and laser heterodyne spectroscopy techniques from balloon and free-flyer platforms, especially in obtaining simultaneous measurements of a large number of source, sink, and intermediate gas molecules in the stratosphere.

At the Langley Research Center, an experiment to measure stratospheric trace constituents by Laser Heterodyne Spectroscopy has been under study for several years, and a summary of sensitivity analyses and supporting laboratory measurements was recently published (ref. 6). The basic model used for the sensitivity studies included: (1) a line-by-line radiative transfer code to calculate transmission of solar radiation through the upper atmosphere in a limb scanning mode, in narrow spectral intervals (40 to 1000 MHz) spanning the infrared from 2.0 to 15.0 μm ; (2) an instrument transfer code to simulate the instrument response function, converting the solar irradiance to digitized voltages as a function of altitude; and (3) an inversion algorithm to convert the digitized voltages to vertical gas concentration profiles. Results were presented for ClO and O_3 for a generalized two gas simultaneous LHS experiment.

In this paper, similar sensitivity analyses are presented for O_3 , ClO , and H_2O_2 in which the instrument transfer function is modeled using a detailed optical receiver design currently under study at the Langley Research Center and using recent laboratory measurements of various subsystems of the LHS instrument. The laboratory measurements have been reported in parallel papers at this conference (refs. 7, 8). The spatial frequency response of an LHS-type

Spacelab instrument has also been reported at this conference (ref. 9), and presents results of tradeoff studies performed to achieve a vertical resolution ≤ 2 km and optimum signal-to-noise ratio for the measurement. Results reported in this paper incorporate the results of these other studies and provide performance criteria required for key components of the LHS experiment including the tunable semiconductor lasers, the photomixers, and the pointing and tracking optics. The optical receiver design and supporting analyses and measurements are used to determine projected accuracies for measurement of the vertical profiles of ClO , O_3 , and H_2O_2 .

INSTRUMENT TRANSFER FUNCTION (ITF)

ITF Model

The model for the ITF used in the LHS sensitivity analyses is shown in figure 1 and accounts for processing of the input solar radiometric signal and the laser local oscillator through the respective transfer optics, photomixer, wideband preamplifier, five intermediate frequency (I.F.) amplifiers, square law detectors and low pass filters.

The power spectral density of the signal and noise for a given I.F. frequency through the I.F. channelization and amplification stages is given as

$$G_{\text{IF}} = e^2 \alpha P_T |H_{\text{PA}}(f_o)|^2 |H_{\text{IF}}(f)|^2 + \frac{1}{2} \frac{k T_{\text{eff}}}{R_m} |H_{\text{PA}}(f)|^2 |H_{\text{IF}}(f)|^2 + e^2 \alpha^2 P_{\text{LO}} [S^*(f + f_o) + S(f - f_o)] |H_o(f)|^2 |H_{\text{PA}}(f)|^2 |H_{\text{IF}}(f)|^2 \quad (1)$$

where $S(f - f_o)$ = the spectral power at the photomixer (ergs/sec-cm^{-1}),
lower sideband

$S^*(f + f_o)$ = the spectral power at the photomixer (ergs/sec-cm^{-1}),
upper sideband

T_{eff} = effective noise temperature of the preamplifier

e = electronic charge

α = η/hf

f = electrical frequency

f_o = local oscillator frequency

η = quantum efficiency

P_T = power required to generate the total shot noise present in the photomixer

R_m = mixer resistance

P_{lo} = local oscillator power

$|H_O(IF)|^2$, $|H_{PA}(IF)|^2$, and $|H_{IF}(IF)|^2$ are the power spectral density transfer functions of the photomixer, preamplifier, and intermediate filter/amplifier filter combination, respectively. In the expression for the spectral power density at the photomixer, $S(f - f_o)$, it is assumed that the heterodyne receiver integrates over a solid angle field of view (Ω) and detector area (A), according to Siegman's antenna theorem (ref. 10). The model takes into account appropriate losses due to polarization effects, and chopping of the solar signal (ref. 11), as well as the appropriate optical losses due to the transfer optics. The final output of the signal after the low pass filter is given by

$$G_{LP}(f) = \int_{-B_{LP}}^{+B_{LP}} G_{SQ}(t) |H_{LP}(f)|^2 df \quad (2)$$

Assuming $|H_{LP}|^2 = 1$ from $-B_{LP} \leq f \leq B_{LP}$, equation (2) becomes

$$G_{LP}(f) = \left[\int_{-\infty}^{\infty} G_{IF}(\mu) d\mu \right]^2 + 2 B_{LP} \left[2 \int_{-\infty}^{\infty} G_{IF}(f - \mu) G_{IF}(\mu) d\mu \right] \quad (3)$$

The heterodyne signal is then

$$S(I_{DC}) = e^2 \alpha^2 P_{lo} 2 \int_{B_{IF}} [S^*(f + f_o) + S(f - f_o)] |H_O(f)|^2 |H_{PA}(f)|^2 df \quad (4)$$

and the rms noise becomes

$$N(I_{rms}) = \left\{ 8 B_{LP} \int_{B_{IF}} \left[(e^2 \alpha^2 P_T + \frac{k T_{eff}}{2 R_m} + (e^2 \alpha^2 P_{lo})(S^* + S) |H_O(f)|^2 |H_{PA}(f)|^2) \right] df \right\}^{1/2} \quad (5)$$

where the IF filter response, $|H_{IF}(f)|^2 = 1$. The first term in equation (5) represents current induced shot noise, the second term describes the thermal noise, and the last term results from the mixing of the noise and signal in the square law device. For the Spacelab experiment, B_{LP} is chosen to filter out as much high frequency noise as possible but still retain sufficient high frequency components of the signal to allow spatial variations in the signal to be measured with a resolution ≤ 2 km. In the model, the dc signal currents are assumed to be digitized with a digitization error of ± 0.1 percent; the dc signal currents are utilized with the inversion algorithm to retrieve gas concentration profiles as a function of altitude (ref. 6).

CONCEPTUAL DESIGN OF A TWO-GAS LHS OPTICAL RECEIVER

Optical Receiver

Figure 2 is a schematic diagram of a two-gas simultaneous LHS optical receiver which is currently being used to provide performance criteria for the ITF model. The optical receiver consists of two optical receiver channels in an integrated package, bore-sighted to observe the same tangent altitude in the stratosphere. Functionally, the optical receiver is composed of three basic subsystems including: (1) the transfer optics to simultaneously irradiate the photomixers with the solar and local oscillator (λ_0) radiation; (2) a wavelength identification and control system to identify the frequency of the λ_0 to the precision required to invert the radiometric signal, and a microprocessor controlled command system to lock the λ_0 wavelength to within 5 MHz (precision) of the center of the absorption line to be measured; and (3) a calibration system with a 1273°K blackbody reference source to provide calibration of the instrument prior to and following measurements performed during the solar occultation measurement period.

The λ_0 transfer optics couple the laser radiation to the optical train of the solar signal through either an f/1 germanium lens, or an f/1.5 parabolic mirror at a 50:50 zinc selenide beam splitter. The semiconductor laser is assumed to be completely polarized along the p-n junction (p-type polarization). The transfer optics for the solar radiation consist of a telescope designed to provide a 3-inch collimated beam as the input to the optical receiver. A reflecting beam divider separates the collimated beam into two beams which serve as the input optical trains to the LHS optical receiver. Each optical train is directed to a diffraction grating, blazed to the wavelength of interest and controlled by a shaft encoder with inputs from a central microprocessor system. The diffracted solar signal is focused on the photomixer through two off axis parabolas (f/3) and combined with the λ_0 radiation at the Zn-Se beam splitter. The angle of the diffraction grating is set to center the bandpass of solar radiation at the wavelength of the absorption line of the target molecule. Laboratory experiments have been performed to demonstrate that a diffraction grating blazed at the appropriate wavelength (for NH_3 , $\nu \approx 965 \text{ cm}^{-1}$, 100 grooves/mm) and a bandpass of 1.5 cm^{-1} at the photomixer, yields a spectral rejection ratio of (1000:1) for solar radiation outside the bandpass.

An important design criterion for the optical receiver shown in figure 2 is the lack of a mode rejection system for radiation from the tunable semiconductor laser. This design is based upon results reported by Jäckel and Guekos (ref. 12) in which the high frequency noise intensity spectra from CW GaAlAs semiconductor lasers were studied, and it was generally observed that a lasing mode which is optically isolated from other axial modes of the laser exhibited much stronger intensity fluctuations when compared to the total output of the semiconductor laser, particularly in the frequency range below the intrinsic natural resonance frequency of the laser. Similar noise measurements have been performed with PbSnSe lasers at the Langley Research Center and have been reported at this conference (ref. 7) confirming these conclusions for Pb-salt lasers. Although all Pb-salt lasers which have been tested do not exhibit these noise properties, the number is sufficiently large with present

state-of-the-art lasers to warrant this design. Therefore, in the design shown in figure 2, the solar signal is filtered to a sufficient extent to require the local oscillator to have a mode spacing $>1.5 \text{ cm}^{-1}$ in order to confine heterodyning with the solar signal to one axial mode.

In designing the optical receiver in which several modes simultaneously irradiate the photomixer, the added shot noise induced by the extraneous axial modes must be considered in the evaluation of the ultimate signal-to-noise ratio as well as excess noise effects (refs. 6, 7). To minimize effects of the added shot noise on the photomixer current, a requirement must be imposed upon the semiconductor laser which takes into account both excess noise and lack of spectral purity. To develop this requirement, a series of semiconductor laser performance criteria have been generated based upon the following analysis. The signal-to-noise ratio for the heterodyne process can be written as

$$S/N = \kappa \frac{P_S P'_{\ell o}}{2e^2 \alpha P'_{\ell c} + \frac{(F-1)kT_o}{R_m} + 2e^2 \alpha N_e} \quad (6)$$

where $P'_{\ell o}$ = local oscillator power per mode within the bandpass of the filtered solar signal

N_e = equivalent excess noise power

κ = a proportionality factor containing various system constants

P_S = blackbody power

F = noise figure

T_o = 273°K reference temperature for noise figure

Neglecting the amplifier noise term (second term in the denominator) this equation can be rewritten as

$$S/N = \kappa' \frac{P_S P'_{\ell o} / P_{\ell o}}{1 + N_e / P_{\ell o}} \quad (7)$$

and $\kappa' = \kappa / 2e^2 \alpha$. If the laser spectral purity is defined as $S_p = P'_{\ell o} / P_{\ell o}$

and the excess noise ratio as $E_{NR} = 1 + N_e/P_{\ell o}$, equation (6) can be written as

$$S/N = \frac{\kappa' S_p}{E_{NR}} = \kappa' F' \quad (8)$$

where $F' = S_p/E_{NR}$ is a figure of merit which defines minimum performance specifications for tunable semiconductor lasers. For a large number of tunable semiconductor lasers which have been tested, the parameter E_{NR} ranged from 1.0 to 1.8 in the quiet regions of the tuning range which for a value of $S_p = 1.0$ results in a typical figure of merit, $F' = 0.55$. This represents a degradation of approximately 0.50 from the theoretically calculated S/N ratio, and is the figure of merit used in the ITF as typical of the expected performance criteria for semiconductor lasers to be used in the optical receiver shown in figure 2. For semiconductor lasers in which the spectral purity (S_p) is < 1.0 , E_{NR} must be less than 1.8 to maintain a degradation of 0.5 (e.g., for $S_p = 0.8$, $E_{NR} = 1.6$; for $S_p = 0.5$, $E_{NR} = 1.0$). Moreover, by measuring the total laser power of the semiconductor laser, the figure of merit F' , and the noise figure of the preamplifier, the expected signal-to-noise ratio for the measurements can be obtained.

In specifying performance criteria for high performance but still practical tunable semiconductor lasers as ℓo 's in the experiment, the following criteria are established at the exit window of the cryogenic cooler:

Total Power $\geq 700 \mu\text{Watts}$ in "P" polarization

Spectral purity ≥ 80 percent in central mode

Mode separation $\geq 1.5 \text{ cm}^{-1}$

$E_{NR} \leq 1.8$

These performance criteria assume the following optical efficiencies in the optical receiver design. Total ℓo radiation losses include the beam splitter (0.5), overfill losses of the ℓo beam image at the photomixer (0.10), and transmission losses of the photomixer window (0.04) for a total transmission efficiency of 43 percent. With 700 μW power at the exit window of the cryogenic cooler, 299 μW of power irradiates the photomixer element which is sufficient to produce 1425 μamps of shot noise current. In the optical train of the solar radiation, optical reflection losses are 0.85 and the loss at the beam splitter is 0.5, for a total transmission efficiency of 0.43. Furthermore, due to the inability to perfectly match the Airey pattern of the solar radiation and the laser ℓo beam, a heterodyne efficiency of 0.84 is assumed in the model. Calculations and trade-off studies performed to optimize the heterodyne efficiency for this design are given in a companion paper to this conference (ref. 9).

Photomixer, Preamplifier, and IF Processing Electronics

The effect of the photomixer, preamplifier, and IF processing electronics on the signal is modeled using power spectral density transfer functions. The

photomixer is modeled using an equivalent electrical circuit (ref. 13) with signal rolloff and shot noise effects included. In the results of the sensitivity analysis reported here, the measured frequency response for photomixer A shown in figure 3 was used from data generated in laboratory experiments (ref. 8). The transfer function was calculated by normalizing curves to a dc value of 1. The shot noise components included in the equivalent circuit are due to the dc current induced by the signal, the local oscillator, and the dark current. In addition, provision was made in the model for a shot noise component induced by excess noise. Noise sources in components following the photomixer are primarily Johnson noise and are combined with the Johnson noise of the photomixer.

The IF processing electronics represent the straightforward application of multipole filters, power distribution amplifiers, and square law RF detection diodes for power detection. The final subsystem in the signal chain is composed of a full-wave synchronous demodulator, running mean integrator and sample and hold circuit. Each element just described was checked to determine fidelity to specification or design and, where necessary, incorporated into the ITF.

Wavelength Identification and Control Subsystem

In figure 2, a wavelength identification and control system is shown in which the laser radiation from the beam splitter is transmitted through an in/out gas cell, a diffraction grating used to isolate the axial mode of the laser which overlaps the absorption line of interest and a 2.5-cm (1-in) diameter in/out Fabry-Perot etalon inserted prior to detection with an Hg-Cd-Te wavelength identification (λ ID) detector. The primary functions of the wavelength ID and control subsystem are to (1) tune the wavelength of the λ_0 to the center of the absorption line to an accuracy between 5 to 20 MHz, depending upon the species of interest, (2) provide command and control logic to lock the wavelength to the required accuracy, and (3) provide command and control logic to the cryogenic cooler and laser wavelength drive to utilize a laser operating point free of "excess-noise" effects.

The wavelength ID and control system has been designed to perform identical wavelength calibration functions as used in high resolution spectroscopic measurements of gas molecules (ref. 14), but the control functions in this experiment will be completely automated and controlled by a central microprocessor system. The wavelength ID system is being assembled to perform the following functions in sequence. First, a coarse wavelength selection of the laser mode is performed using a servo-driven grating monochromator at the same time that the solar isolation grating is tuned to the center of the atmospheric absorption line to be measured. A microprocessor system sets the cooler temperature and ramps the injection current to determine operational limits that locate the laser mode within the bandpass of the wavelength identification monochromator. The spectral response of the λ ID detector is stored by the microprocessor as well as the threshold current I_1 and the current at mode power cutoff I_2 . Secondly, the gas cell is driven into the laser beam and the laser signal is ramped again. The reference gas in the cell must be a stable gas with at least two well-known vibration-rotation lines lying on

either side of the atmospheric absorption line to be measured. Suitable reference gases for the 7.5 to 13.0 μm region include OCS, SO_2 , and NH_3 . The spectral response of the λID detector is again stored by the microprocessor and the two response functions ratioed to obtain the normalized absorption spectra of the reference gas. Third, the Fabry-Perot etalon with free spectral range of 500 MHz is inserted, with the reference gas cell removed, to generate interference fringes to establish the tuning rate for the laser mode. Following this operational step, the microprocessor has sufficient information to set the current of the laser to lock the wavelength near the center of the atmosphere absorption line.

Fine tuning of the λ_0 wavelength is achieved by microprocessor interrogation of the five I.F. signal channels to insure during the first solar occultation measurement period that the spectral distribution of the five I.F. channels conforms to expected values. The microprocessor system will be designed to provide this fine tuning capability through control of the temperature and injection current of the semiconductor laser. In the event that the expected absorption line is not identified in the atmosphere, the microprocessor control unit will be placed in a scanning mode of operation to generate absorption spectra of the atmosphere over the tuning range of the TDL mode.

Calibration and Alignment Subsystem

The calibration subsystem shown in figure 2 consists of an 1273°K black-body source and a collimator assembly unit designed to completely retrace the solar signal through the two channels of the LHS optical receiver, and provides for injection of a highly stable, ruggedized CO_2 laser as a reference λ_0 signal to validate performance of the photomixer and the I.F. processing electronics. The calibration system will be used for ground calibration and alignment. A He-Ne laser with a divided two-beam output will be used for ground set-up and alignment (detail not shown on the figure).

LABORATORY MEASUREMENTS

A series of laboratory measurements has been performed to substantiate some of the basic design criteria shown in figure 2 with a laboratory 1-channel LHS instrument. A schematic of the laboratory measurement system is shown in figure 4. The instrument system consisted of three radiation sources including a 1273° K blackbody source, a highly stable CO_2 laser and a tunable semiconductor laser mounted in a standard cryogenic mechanical refrigerator. Experiments were performed in which the blackbody source could be heterodyned with either the CO_2 or semiconductor laser. Similarly, heterodyning experiments have been performed with the CO_2 and TDL lasers. Two photomixers were used in the experiment including a commercial photomixer with a measured effective heterodyne quantum efficiency (η) of 12 percent from 5 to 115 MHz, and a research photomixer with a measured η of 38 percent. The research photomixer was provided by the MIT Lincoln Laboratories, D. L. Spears. Preamplifier noise figures for both photomixers were 2 dB. In general, the laboratory experiments addressed three basic questions:

- (1) The signal-to-noise ratio obtained with the tunable semiconductor

lasers when compared to the signal-to-noise measurements obtained with a stable CO₂ laser.

(2) The phenomena of "excess noise" and expected degradations for typical tunable semiconductor lasers either procured under contract, obtained from NASA sponsored research programs,¹ or obtained on consignment from the General Motors Research Laboratories.²

(3) The relative magnitude of the RF noise induced by photomixing a multiplicity of axial modes from the tunable semiconductor lasers and the blackbody source and the RF noise induced by a frequency isolated mode from the semiconductor lasers.

In order to determine the projected sensitivity of the LHS experiment to measure trace gas constituents in the stratosphere, laboratory measurements have been performed to compare signal-to-noise ratios with selected semiconductor lasers and highly stable CO₂ laser local oscillators. Figure 5 shows comparative results of the signal-to-noise ratio for measurements with CO₂ lasers (solid lines) and a semiconductor laser (squares). This particular semiconductor laser exhibited nearly single mode output over the current tuning range for a fixed operating temperature. The measured signals for the CO₂ laser measurements were compared to a theoretically derived curve for a 1273° K blackbody, with appropriate losses due to the transfer optics. Comparison between theory and experiment is shown in the figure for the two photomixers. The experiments performed with the semiconductor laser generated a maximum heterodyne detector photocurrent of 400 microamperes ($\eta_{\text{HET}} = 38$ percent). The higher signal-to-noise ratio for the semiconductor laser results from minor alignment differences and experimental scatter. The important point of these results is that semiconductor lasers, when properly selected for quiet operation, can be as effective as CO₂ lasers in providing signal-to-noise at equal power.

Utilizing the experimentally measured value of 1273° K, the experimental curve was extrapolated to a solar source temperature of 5500° K for an assumed effective heterodyne quantum efficiency of 25 percent and other parameters listed in the figure. These results demonstrate that for semiconductor lasers operating with single mode output, signal-to-noise ratios in excess of 1000 can be achieved using heterodyne detection without requiring the heterodyne photocurrent to be in the shot noise limited regime.

In general, tunable semiconductor lasers operating in a single mode over desired wavelengths for the LHS experiment are difficult to obtain with current technology.

¹Research lasers were obtained from Laser Analytics Incorporated under contract NAS1-15190. Partial funding for this contract was provided by the Los Alamos Laboratories, technical monitor Dr. H. Flicker.

²A number of research lasers were obtained from the General Motors Research Laboratories on consignment for test and evaluation in the heterodyne mode. Technical collaborators included Dr. John Hill and Dr. Wayne Lo.

The "excess noise" effects in certain regions have been discussed by Ku and Spears (ref. 15). In a companion paper to this conference, "excess noise" measurements and mechanisms have been discussed (ref. 7). The LHS design shown in figure 2 assumes that semiconductor lasers will be multimode and will operate in the quiet regions of the TDL output. Control of "excess noise" regions by controlling operating temperature and injection current has been discussed previously (ref. 6). In figure 6, comparative measurements of the RF noise power in dB as a function of frequency from 5 to 500 MHz is shown for a typical tunable semiconductor laser and a CO₂ laser local oscillator for a signal induced photocurrent of 400 microamperes. For reference, the preamplifier and dark current noise level for the photomixer and IF processing electronics is shown. As noted earlier, it has been observed that in "quiet" regions of the laser tuning range, the RF noise power can still be more severe for semiconductor lasers than for CO₂ lasers. Measurements were performed using a semiconductor laser in which greater than 80 percent power was in the central mode and the remaining power distributed over three other axial modes. Typical values of the degree to which TDL RF power exceeds that from equivalent CO₂ laser induced RF power are on the order of 2 dB. This residual noise has been used in the ITF in computing the uncertainty in the retrieved concentrations of various trace molecules in the stratosphere using the LHS experiment. For completeness, the measured dependence of the photomixer efficiency as a function of frequency for two photomixers studied in our laboratory is shown in figure 3. In the ITF, photomixer A has been modeled as the expected performance criteria for the LHS experiment.

LHS RETRIEVED CONCENTRATION PROFILES (SIMULATED)

Figure 7 shows results of the LHS sensitivity studies for measuring stratospheric O₃ from Spacelab altitudes, using the ITF function discussed in section 2. For this simulation, an λ_0 wavelength at 1129.4420 cm⁻¹ was selected to correspond to the peak of a relatively intense O₃ molecular transition lying within an atmosphere window. Five I.F. channels were selected to detect radiances within the O₃ line in order to optimize the sensitivity of the measurement over the entire altitude range. The position of the I.F. channel centers relative to the λ_0 wavelength are listed as DNU (cm⁻¹); the bandpass of each channel is listed as BETA (MHz). An integration time, TAU (sec), of 100 ms is used for each channel. For each tangent altitude, two channels are used to invert the radiance data. One channel lies within the O₃ absorption line, and the other channel lies outside the absorption line to account for fluctuations in the background radiance and for continuum absorption effects. For upper altitudes where O₃ attenuation is relatively low, channels near line center are used; for lower altitudes where the O₃ attenuation is high, channels in the wing of the O₃ line are used. The channels selected for various altitudes are listed to the right of figure 7.

In figure 7, the two solid profiles represent the initial and mean retrieved O₃ profile. Error bars indicate ± 1 standard deviation of 20 sample retrievals generated by perturbing the simulated instrument current profile with a random Gaussian noise source to simulate random instrument errors. The error bars are an estimate of the uncertainty in the LHS measurement of ozone at various tangent altitudes. Figure 8 shows the fractional error of the

mean mixing ratio for the simulated measurements as a function of altitude for the profile in figure 7. Measurement uncertainties ≤ 7 percent can be achieved for altitudes ≥ 20 km. For altitudes below 20 km, the fractional error increases due to the relatively small O_3 concentrations in this altitude range and the large attenuation by O_3 in the outer shells. To obtain a fractional error of the mean ≤ 7 percent below 20 km requires the use of a second local oscillator tuned to a weaker O_3 transition in order to increase the solar radiance incident upon the Hg-Cd-Te photomixer. Use of two λ_o 's for the ozone measurement provide measurement uncertainties ≤ 7 percent over the total profile.

Figures 9 and 10 show similar results for ClO for an λ_o wavelength of 856.499 cm^{-1} (ref. 16). In figure 9, two retrieved profiles are shown for the same initial profile. In figure 9(a), the conventional inversion algorithm was used as described in (ref. 6). In figure 9(b), a data smoothing cubic spline routine was used to reduce the random errors associated with the retrieved profile. The results show a significant improvement in the retrieved profile where error bars representing ± 1 standard deviation are significantly reduced (factor of 6). Although use of the cubic spline technique significantly reduces the error bars, use of the technique will reduce the capability of the experiment to measure spatial variability in the vertical profile for a spatial extent ≤ 2 km. Figure 10, which shows the fractional error of the mean mixing ratio for ClO , corresponds to the unsmoothed retrieved profiles.

Figures 11 and 12 show similar results for H_2O_2 for an λ_o wavelength of 1251.2563 cm^{-1} , again for an unsmoothed inversion algorithm.³ For these simulations, it has been assumed that atmospheric pressure at the tangent altitude can be measured to an accuracy of ± 3 percent (ref. 17), and the temperature profile is determined to an accuracy of ± 3 K. Results of the studies performed on O_3 , ClO , and H_2O_2 with the current ITF indicate the following: O_3 retrievals over the vertical profile from 20 to 50 km can be obtained with accuracies exceeding 93 percent without data smoothing and with accuracies approaching 96 percent with data smoothing; for ClO , accuracies ≥ 50 percent can be achieved over 20 to 40 km without data smoothing and ≥ 85 percent with data smoothing; for H_2O_2 , accuracies ≥ 90 and 80 percent can be achieved over 20 to 40 km with and without use of data smoothing techniques, respectively. These sensitivity analyses are currently being conducted for other scientifically interesting gases important to understanding the chemistry of the stratosphere, but depend upon the generation of high resolution spectroscopic parameters including line intensity, line position, and line shapes.

³ H_2O_2 spectra were provided by Dr. A. Goldman, University of Denver ($\Delta\nu \leq 0.02\text{ cm}^{-1}$).

REFERENCES

1. Farmer, C. B.: Infrared Measurements of Stratospheric Composition. Can. J. Chem., 52, 1974, pp. 1544-1559.
2. Russell, J. M., III; Park, J. H.; and Drayson, S. R.: Global Monitoring of Stratospheric Halogen Compounds from a Satellite Using Gas Filter Spectroscopy in the Solar Occultation Mode. Appl. Opt., 16, March 1977, pp. 607-612.
3. Upper Atmospheric Research Satellite Program, Final Report of the Science Working Group, Jet Propulsion Laboratory Publication 78-54, Jet Propulsion Laboratory, Pasadena, CA, June 15, 1978.
4. Chlorofluoromethanes and the Stratosphere. NASA Reference Publication 1010, R. D. Hudson, Editor, Goddard Space Flight Center, Greenbelt, Maryland, 1977, 265 pp.
5. High Resolution Infrared Spectroscopy Techniques for Upper Atmospheric Measurements. Edited by D. G. Murcray and J. M. Alvarez, NASA CP-2134, 1980.
6. Allario, F.; Katzberg, S. J.; Hoell, J. M.; and Larsen, J. C.: An Experiment Concept to Measure Stratospheric Trace Constituents by Laser Heterodyne Spectroscopy. Accepted for Publication in Applied Physics, 1980, Springer Verlag.
7. Harward, Charles N.; and Sidney, Barry D.: Excess Noise in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ Semiconductor Lasers. Heterodyne Systems and Technology, NASA CP-2138, 1980. (Paper 10 of this compilation.)
8. Kowitz, Herbert R.: Comparative Performance of HgCdTe Photodiodes for Heterodyne Application. Heterodyne Systems and Technology, NASA CP-2138, 1980. (Paper 22 of this compilation.)
9. Fales, Carl L.; and Robinson, Don M.: Spatial Frequency Response of an Optical Heterodyne Receiver. Heterodyne Systems and Technology, NASA CP-2138, 1980. (Paper 35 of this compilation.)
10. Siegman, A. E.: The Antenna Properties of Optical Heterodyne Receivers. Proc. IEEE, 54, 1966, pp. 1350-1356.
11. Abbas, M. M.; Mumma, M. J.; Kostiuk, T.; and Buhl, D.: Sensitivity Limits of an Infrared Heterodyne Spectrometer for Astrophysical Applications. Appl. Opt., 15, Feb. 1976, pp. 427-436.
12. Jackel, H. and Guekos, G.: High Frequency Intensity Noise Spectra of Axial Mode Groups in the Radiation from CW GaAlAs Diode Lasers. Opt. and Quant. Electr. 9, 1977, pp. 233-239.

13. Peyton, B. J.; Dinardo, A. J.; Cohen, S. C.; McElroy, J. H.; and Coates, R. J.: An Infrared Heterodyne Radiometer for High-Resolution Measurements of Solar Radiation and Atmospheric Transmission. IEEE J1. Quant. Elect., QE-11, August 1975, pp. 569-574.
14. Brockman, P.; Bair, C. H.; and Allario, F.: High Resolution Spectral Measurements of the HNO_3 11.3 μm Band Using Tunable Diode Lasers. Appl. Opt., 17, Jan. 1978, pp. 91-100.
15. Ku, R. T. and Spears, D. L.: High Sensitivity Heterodyne Radiometer Using a Tunable - Diode - Laser Local Oscillator. Opt. Lett., 1977, pp. 84-86.
16. Rogowski, R. S.; Bair, C. H.; Wade, W. R.; Hoell, J. M.; and Copeland, G. E.: Infrared Vibration - Rotation Spectra of the $\text{C}\ell\text{O}$ Radical Using Tunable Diode Laser Spectroscopy. Appl. Opt., 17, May 1978, pp. 1301-1302.
17. Park, J. H.; Russell, J. M., III; and Drayson, S. R.: Pressure Sensing of the Atmosphere By Solar Occultation Using Broadband CO_2 Absorption. Appl. Opt., 18, June 1979, pp. 1950-1954.

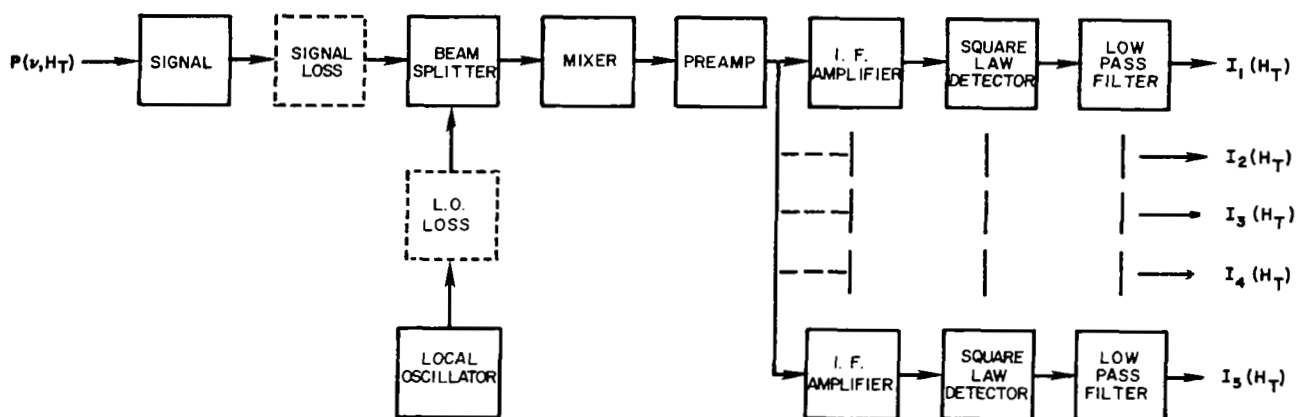


Figure 1.- LHS instrument transfer function model.

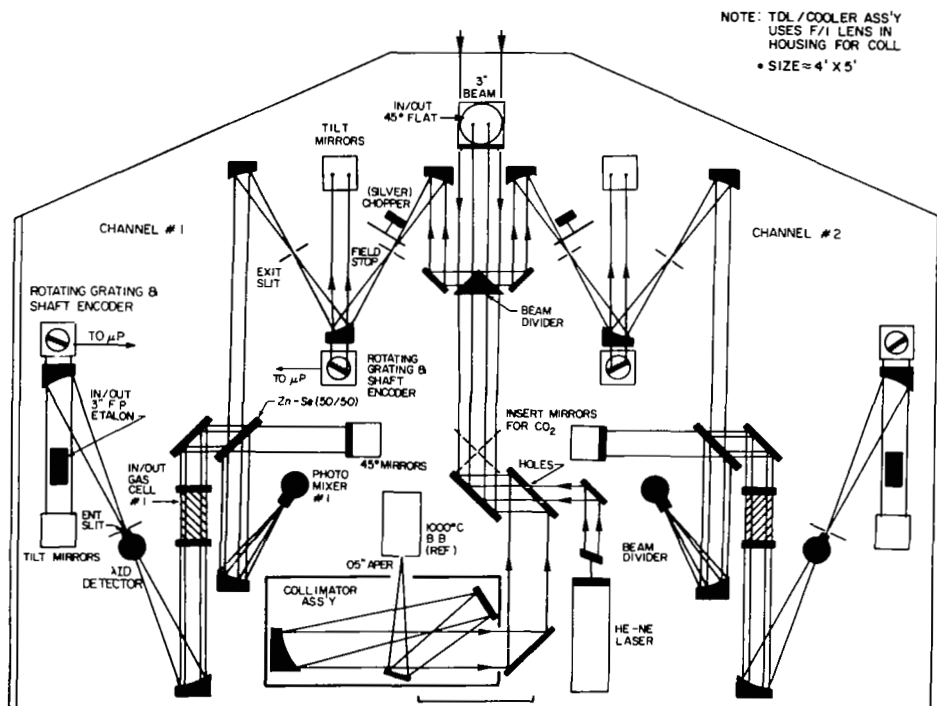


Figure 2.- Schematic of two-gas LHS Spacelab optical receiver.
(1 in. = 2.54 cm; 1 ft. = 0.305 m.)

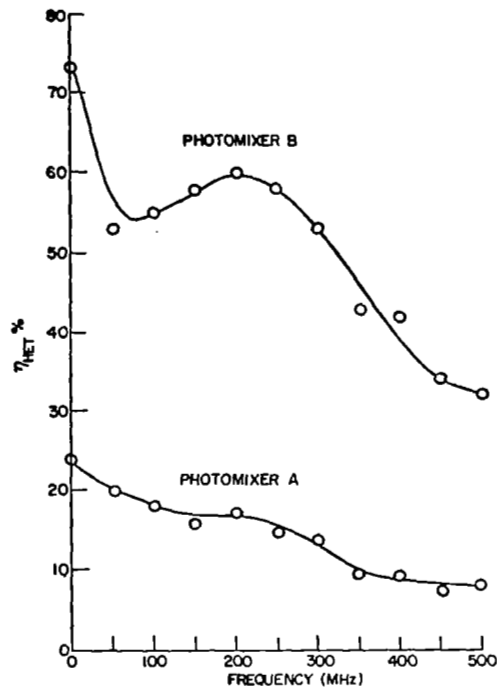


Figure 3.- Measured photomixer quantum efficiency vs frequency. Photomixer A - research photomixer; photomixer B - commercial device.

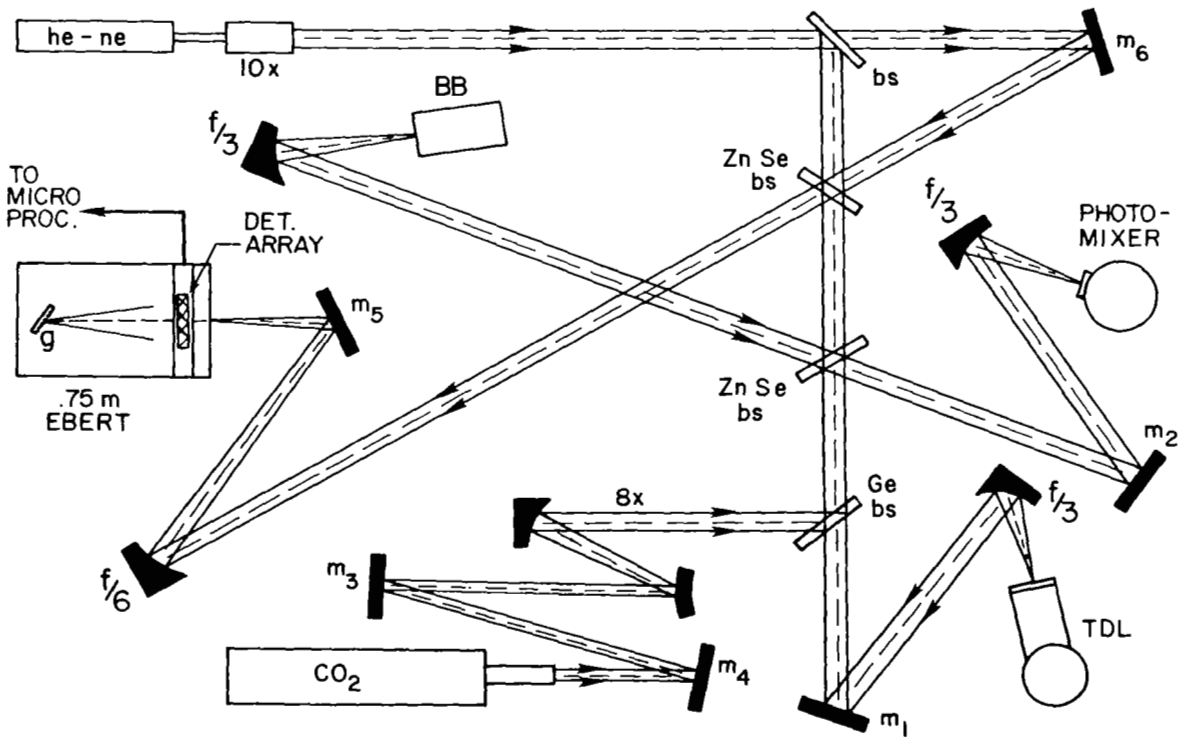


Figure 4.- Single channel LHS laboratory model.

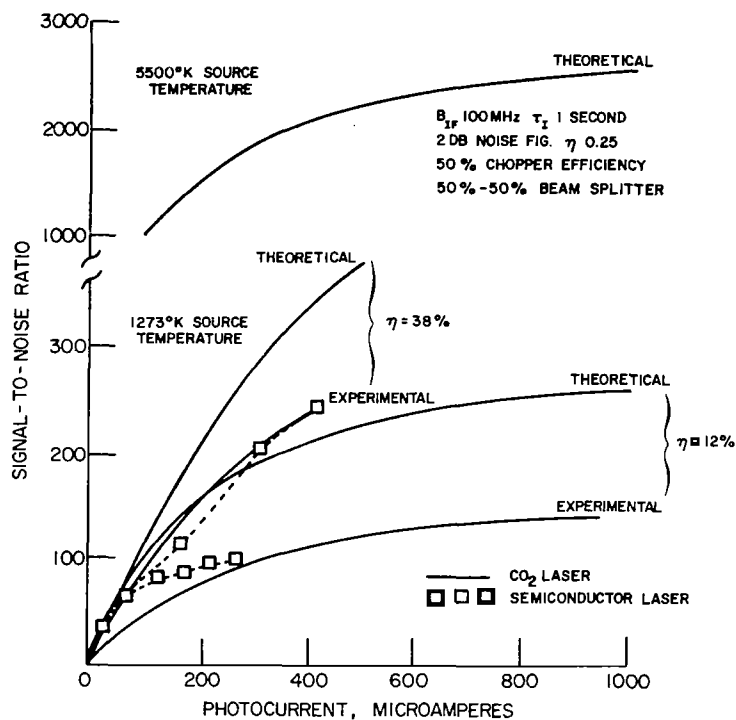


Figure 5.- Comparative signal-to-noise measurements for CO₂ and semiconductor lasers.

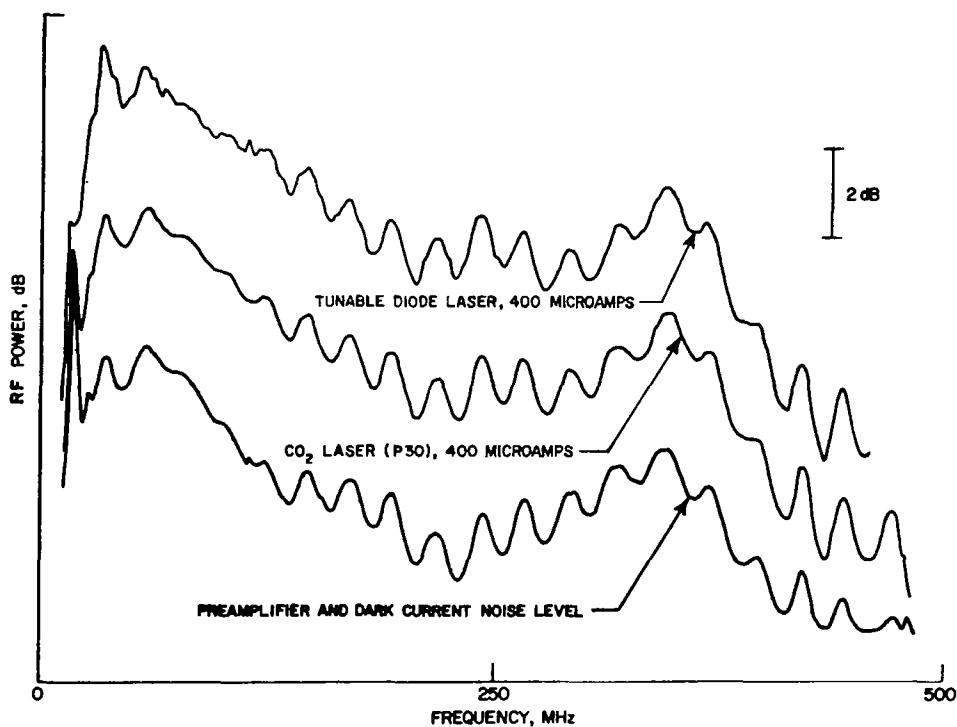


Figure 6.- Comparative RF noise measurements for CO₂ and semiconductor lasers.

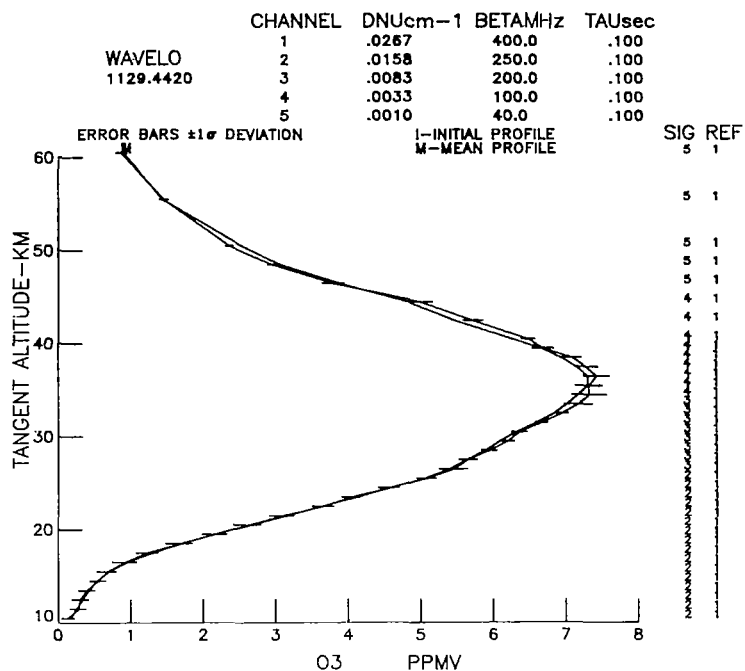


Figure 7.- LHS simulations for stratospheric ozone.

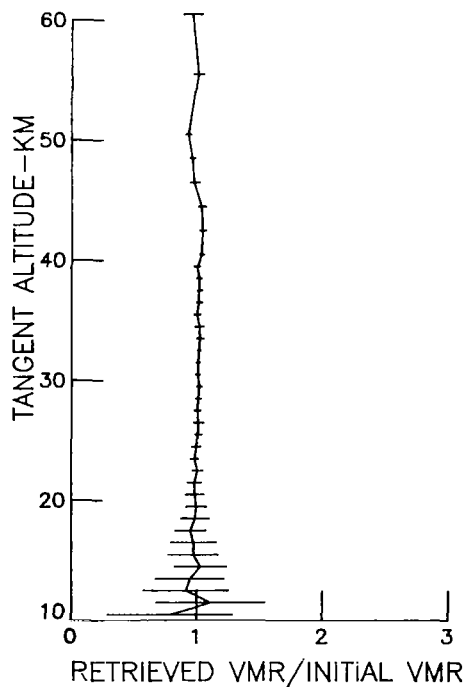
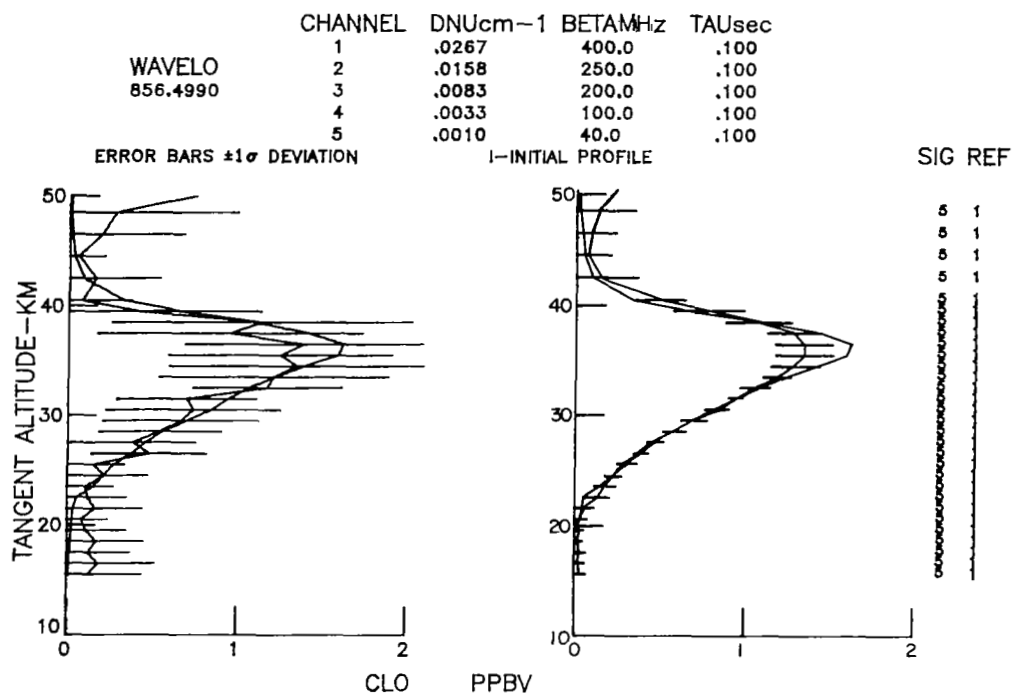


Figure 8.- Fractional error of the mean mixing ratio for O_3 . Error bars $\pm 1\sigma$ deviation.



(a).- Without data smoothing
(cubic spline technique).

(b).- With data smoothing
(cubic spline technique).

Figure 9.- LHS simulations for stratospheric chlorine monoxide.

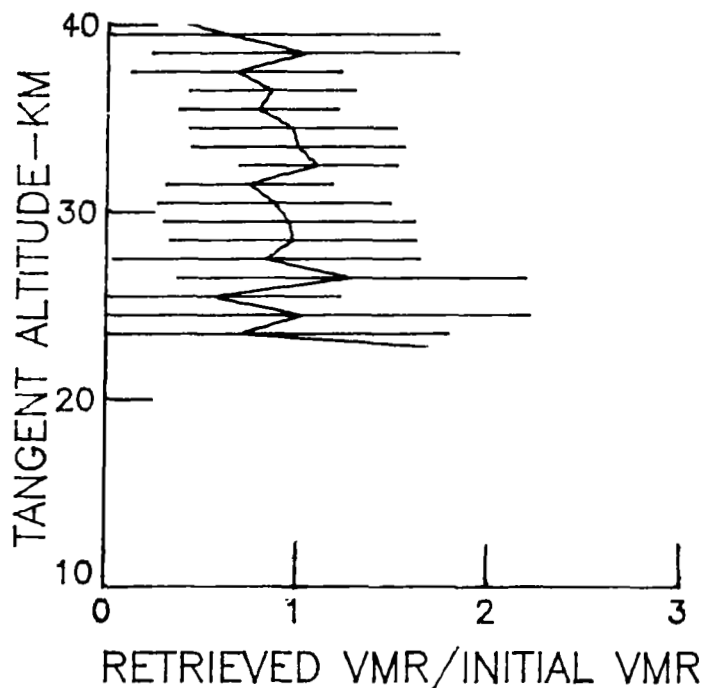


Figure 10.- Fractional error of the mean mixing ratio for ClO. Error bars $\pm 1\sigma$ deviation.

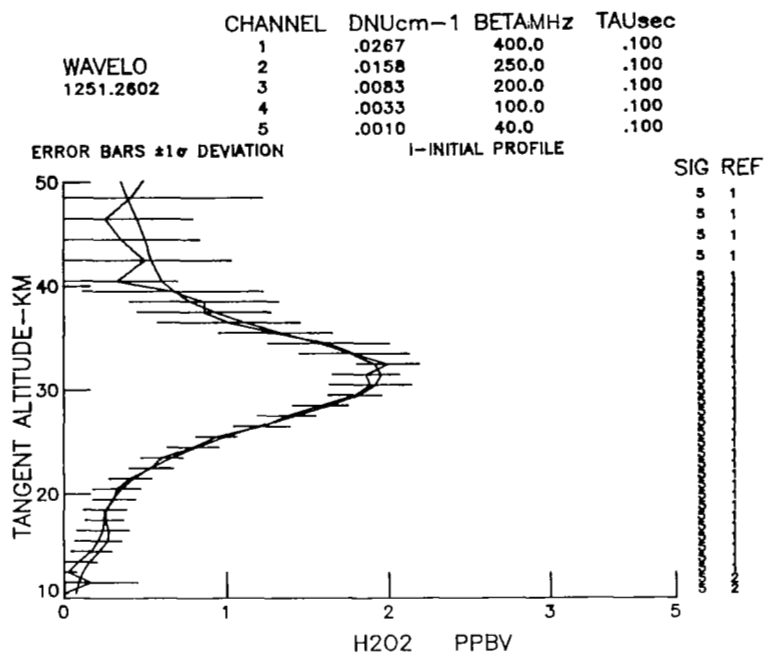


Figure 11.- LHS simulation for stratospheric hydrogen peroxide.

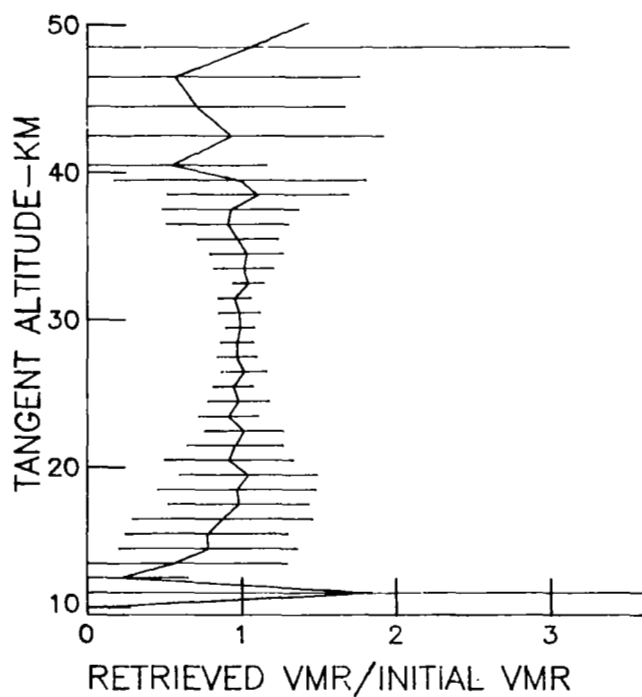


Figure 12.- Fractional error of the mean mixing ratio for H_2O_2 . Error bars $\pm 1\sigma$ deviation.

A 163 MICRON LASER HETERODYNE RADIOMETER FOR OH:

PROGRESS REPORT

H. M. Pickett and T. L. Boyd
Jet Propulsion Laboratory

ABSTRACT

We are currently developing a 163 micron (1.836 THz) radiometer for airplane and/or balloon platforms.

The laser local oscillator is a CO₂ pumped methanol laser operating at frequency which is ~ 1 GHz from the $J = 3/2 - 1/2$ transition of OH. The laser will be used directly as a local oscillator or will be translated in frequency to closer coincidence with the OH emission, depending on achieved detector I.F. bandwidth. Frequency translation techniques which will be described are diode mixing and a new method of single sideband generation using an external Stark modulated gas cell.

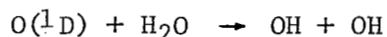
The photoconductive mixer which will be used is a strained Ge crystal, doped with Ga, originally used as an incoherent detector. The uniaxial strain on the Ga doped Ge crystal shifts the threshold for photoconduction from 100 cm⁻¹ to frequencies as low as 50 cm⁻¹. These detectors are currently being characterized as mixers in the laboratory. Of particular interest are the effect of local oscillator power and strain on I.F. bandwidth, detector impedance and conversion loss. Preliminary results of these tests will be described and compared with theoretical expectations.

INTRODUCTION

The hydroxyl radical (OH) is recognized as a keystone species for the understanding of ozone destruction in the stratosphere. As Wofsy has stated (ref. 1),

"The OH radicals.....drive most of the photochemistry of the stratosphere. They play a pivotal role in the partitioning of NO_x and ClO_x among species which are relatively reactive or stable toward O₃ and O."

The OH radical plays a catalytic role in the HO_x cycle. In addition, increased OH would reduce the importance of the NO_x cycle but would increase the importance of the ClO_x cycle (ref. 2). The major source of OH is the reaction of O(1D) with water:



In the region below 80 km, the main source of $\text{O}(^1\text{D})$ is the photolysis of ozone (ref. 3). As a consequence, OH concentration is a maximum at ~ 50 km. The region near 50 km is important for checking on the quantity of OH produced, while the region of 20-40 km is important for determining the relative roles of the chlorine and nitrogen cycles.

The OH submillimeter wavelength rotational lines provide an excellent probe of OH, particularly when remote sensing emission measurements are used. Figure 1 shows the emission of OH calculated for a balloon platform looking horizontally at an altitude of 40 km. Most of the emission shown comes from higher altitude. Calculated emission from an airplane flying at 12 km would be quite similar. Emission from lower altitudes is broader because of increased air pressure and is expected to be weaker because of decreased OH abundance. This particular OH transition is the lowest frequency set of rotational lines and happens to be quite close to a methanol laser line. In figure 2 it can be seen that several of the lines are within 1 GHz of the laser transition. This fortunate coincidence forms the basis for our 1.8 THz radiometer. It might be noted that these lines have recently been observed in the stratosphere by Kendall and Clark (ref. 4) using a Michelson interferometer on a balloon platform. However, for detailed altitude distributions of OH, a heterodyne system will be needed.

LOCAL OSCILLATOR

The block diagram of our radiometer is shown in figure 3. It consists of the usual combination of mixer, local oscillator, and filter bank. If the mixer has a bandwidth in excess of 1 GHz, the laser can be used directly as the local oscillator. However, it is quite likely that the bandwidth of our mixer will be smaller than 1 GHz and a scheme for frequency shifting the laser is needed. The scheme we are planning involves a modulation of the laser followed by filtering of the lower modulation sideband. This modulator and sideband filter are therefore part of the local oscillator system. A number of groups have been active in generating sidebands using an R.F. biased diode (e.g. ref. 5). The efficiency of sideband generation using this method is not large but the offsets achievable are limited only by the optical coupling structure for the diode.

An alternative modulation method we have been investigating involves modulation of absorption in an external gas cell. Surprisingly such a technique can generate in a single sideband mode with up to 10% efficiency. The method is best illustrated by a model experiment we performed at 104 GHz using a methyl fluoride absorption (ref. 6.). The block diagram of the spectrometer system used is shown in figure 4. There is a source klystron locked to 102 GHz and a local oscillator klystron locked to a frequency

which is exactly 1140 GHz below the source klystron. The mixer and spectrum analyzer combine to make a very high resolution radiometer for power leaving the sample cell. The methyl fluoride pressure in the sample cell was ~ 80 mtorr and the 10.8 MHz electric field had an amplitude of 40 V/cm. This field modulates the absorption frequency by 14 MHz. The modulation frequency was significantly larger than the 1.5 MHz width of the methyl fluoride absorption. The results of this experiment are shown in figure 5. Trace B, taken when the input signal was on resonance with the absorption shows the sideband pattern expected for amplitude modulation. When the input is above or below resonance, the effect of the modulated absorption is to put enhanced power in the sideband located on the other side of the absorption center. In this experiment, the absorptivity of the sample was small and the sideband level only reached the 5 db level. However, for stronger absorption, the theoretical model of the experiment predicts that the sideband level can approach 10 db.

THE MIXER

Recently we have been centering most of our attention on the mixer, since the mixer characteristics drive many of the decisions about local oscillator requirements and achievable signal to noise figure. The mixer which we are testing is a strained Gallium doped Germanium crystal which is used in a photoconductive mode. Incoherent detectors of this type have been described (ref. 7), but their mixer properties have not been evaluated. We are currently evaluating the bandwidth of several mixer crystals at 96μ , 118μ and 163μ in the stressed and unstressed conditions. At this point we can only report that the bandwidth in all cases is in excess of 5 MHz based on time response tests. The material used is supplied to us by Santa Barbara Research Corp. and is only lightly compensated. The crystals are in the form of cubes 3mm on a side and are doped to a level of $\sim 10^{14}/\text{cm}^3$. Measurements of I.F. noise bandwidth by mixing a black body with the laser are currently under way.

We have also measured the D.C. photo-resistance at laser power levels in an attempt to determine the maximum achievable conversion gain of the mixer. An example of such a photo-resistance curve is shown in figure 6. As can be seen, the curve fits a simple power saturation law. This saturation behavior is reasonable if the negatively ionized acceptor concentration is not a function of the power level. At higher power levels the ionized acceptor concentration should increase due to the production of photo-holes significantly in excess of the dark ionized acceptor level. This increased ionized acceptor level will 1) increase the recombination rate and 2) decrease the mobility. The first effect occurs because the recombination rate is proportional to both ionized acceptor concentration and hole concentration, and the second effect occurs because of increased scattering by ionized acceptors relative to neutral acceptors. Both effects will slow the saturation rate with increased power, and such a trend can be seen at the higher power levels in figure 6. The conversion gain is obtainable from the

photoresistance by

$$G = 2 \left(\frac{dR}{dP} \right)^2 I_b^2 R^{-1} P, \quad (1)$$

in which I_b is the bias current. The maximum gain then appears at $P = P_0/2$, with

$$G_{\max} = \left(I_b^2 R_0 / P_0 \right) \times 0.296 \quad (2)$$

in which R_0 is the dark resistance. It is possible to make $I_b^2 R_0$ on the order of P_0 , but higher bias levels suffer from the severely non-linear I-V characteristics of the crystals (see ref. 7). To achieve the high gain of eq. (2), the I.F. amplifier has to be matched to an impedance of $\sim 28 \text{ k}\Omega$. This is a problem with conventional $50 \text{ }\Omega$ amplifiers but is quite well suited to FET input amplifiers at moderate I.F. frequencies.

REFERENCES

1. Wofsy, S. C.: Temporal and Latitudinal Variations of Stratospheric Trace Gases: A Critical Comparison Between Theory and Experiment. J. Geophys. Res., vol. 83, 1978, pp.364-378.
2. Crutzen, P. J.; Isaksen, I.S.A.: and McAfee, J.R.: Impact of the Chlorocarbon Industry on the Ozone Layer. J. Geophys. Res., vol. 83, 1978, pp. 345-363.
3. Vlasov, M. N.: Photochemistry of Excited Species, J. Atm. Terr. Phys., vol. 38, 1976, pp. 807-820.
4. Kendall, D. J. W.; and Clark, T. A.: Stratospheric Observation of Far I.R. Pure Rotational Lines of Hydroxyl, Nature, vol. 283, 1980, pp. 57-58.
5. Bicanic, D. D.; Zuidberg, B. F. J.; and Dymanus, A.: Generation of Continuously Tunable Laser Sidebands in the Submillimeter Region, Appl. Phys. Lett., vol. 32, 1978, pp. 367-369.
6. Pickett, H. M.: The Physical Basis for Absorption of Light, Nature, vol. 279, 1979, pp. 224-225.
7. Kayanskii, A. G.; Richards, R. L.; and Haller, E. E.: Far Infrared Photoconductivity of Uniaxially Stressed Germanium, Appl. Phys. Lett., vol. 31, 1977, pp. 496-497.

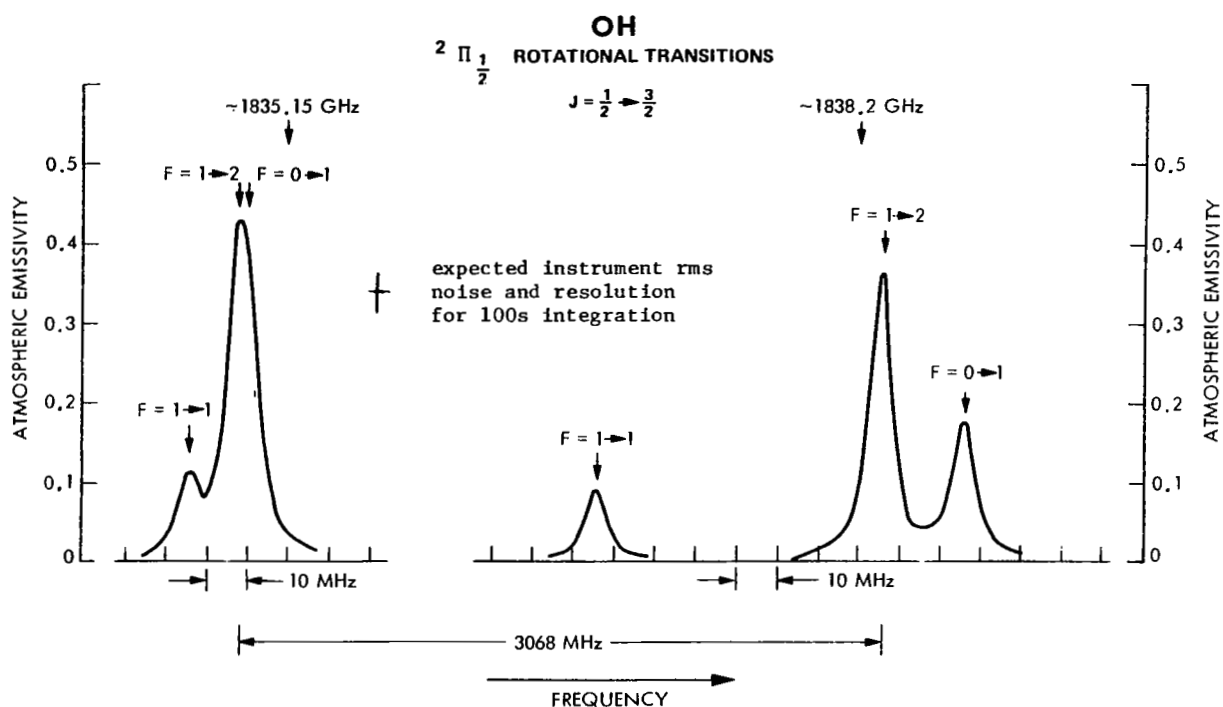


Figure 1.- Calculated emission by OH in Earth's upper atmosphere.

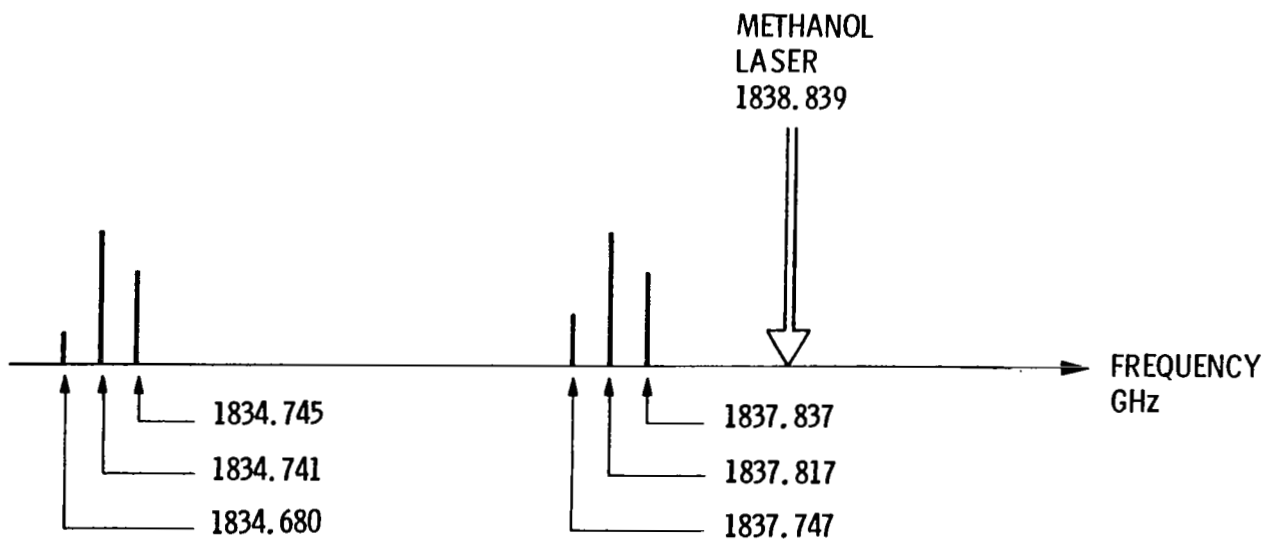


Figure 2.- OH spectrum near 1800 GHz.

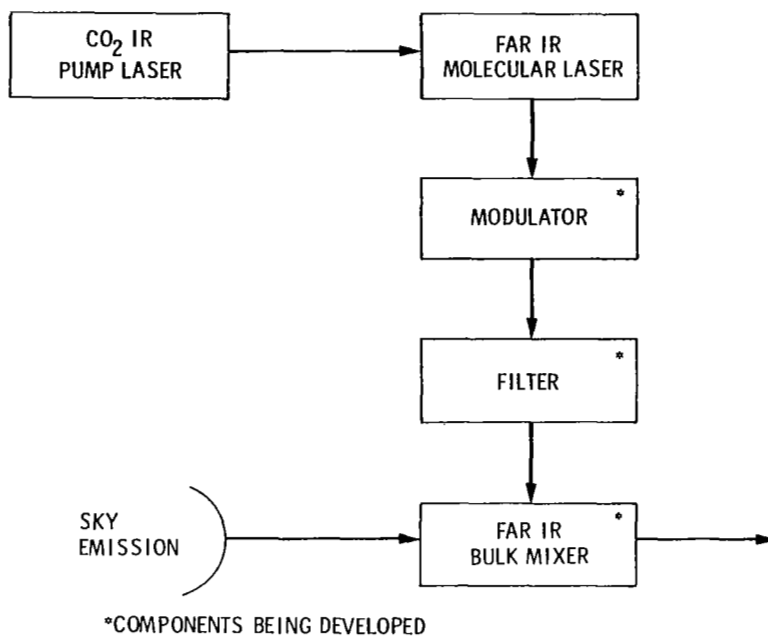


Figure 3.- 1800 GHz OH radiometer.

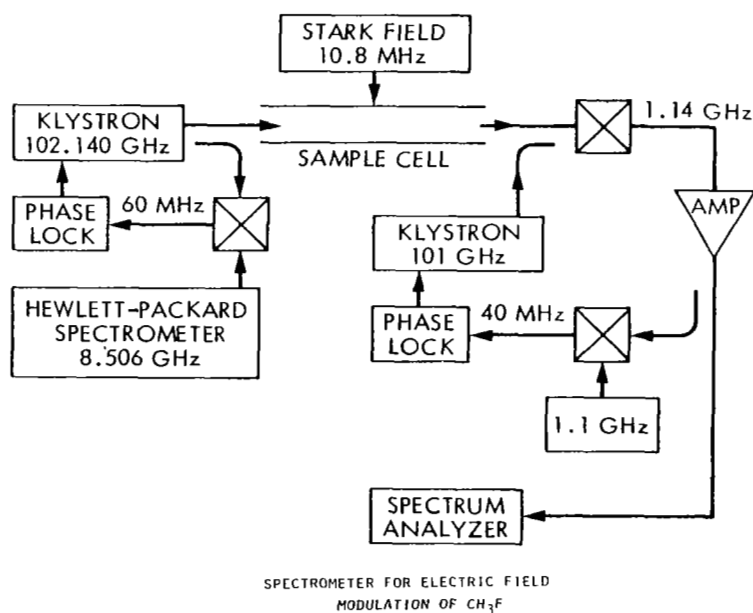
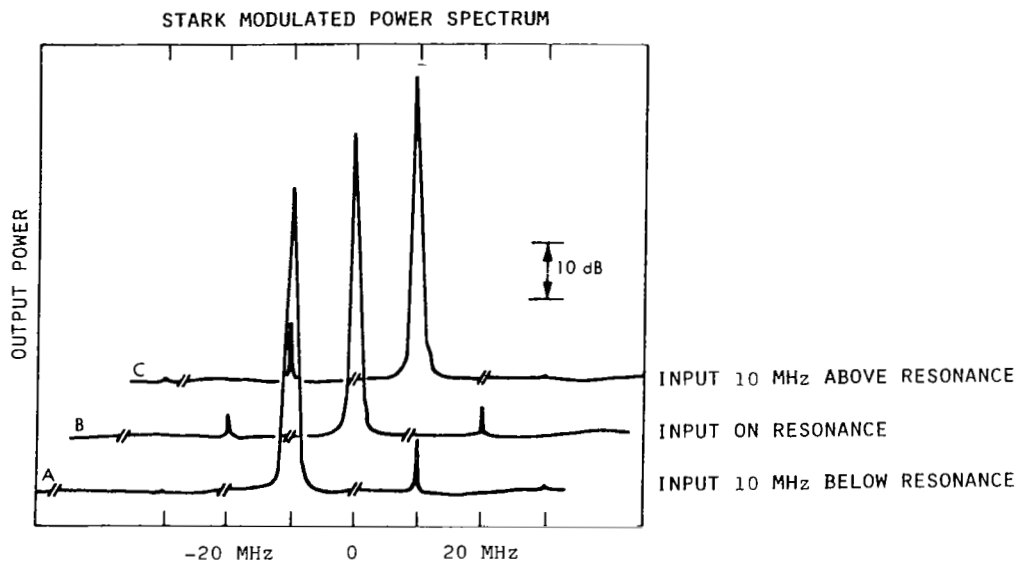


Figure 4.- Laboratory test of modulation scheme.



FREQUENCY OFFSET FROM MOLECULAR RESONANCE OF CH_3F AT 102.140 GHz

Figure 5.- New insights into the physics of light absorption. Measured asymmetry in modulated power requires light absorption to be treated as stimulated emission, not as amplitude absorption.

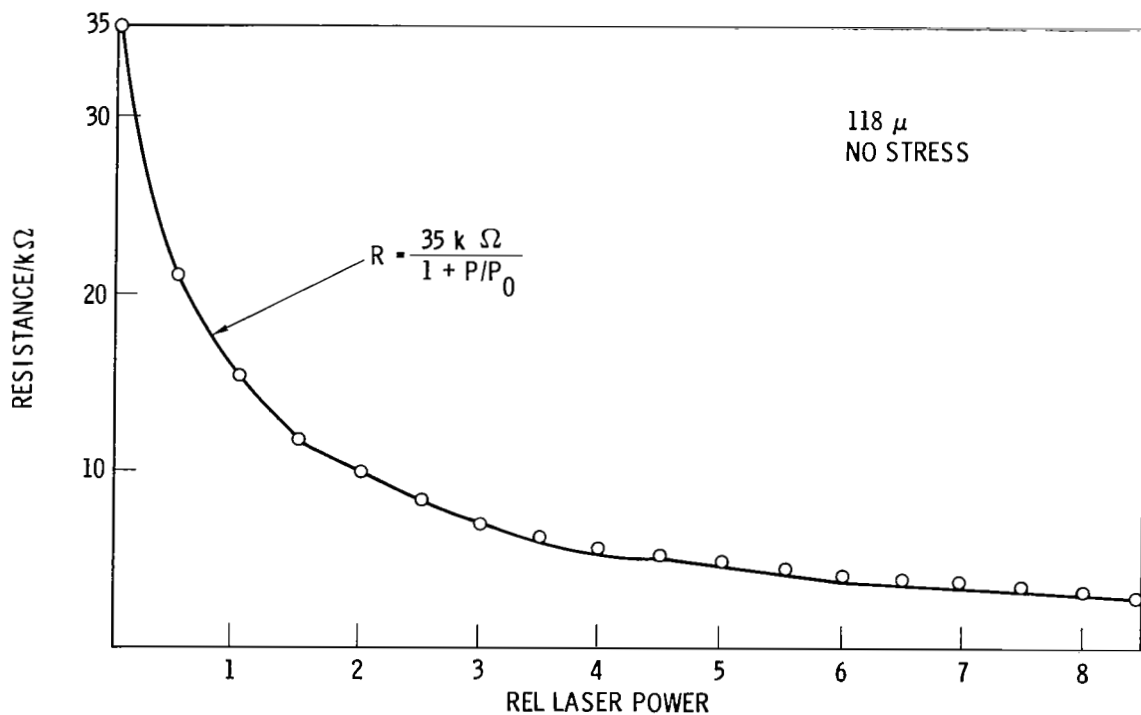


Figure 6.- Photoresistance.

HETERODYNE DETECTION OF THE 752.033-GHz H_2O

ROTATIONAL ABSORPTION LINE

G. F. Dionne, J. F. Fitzgerald, T-S. Chang, M. M. Litvak,* and H. R. Fetterman
Lincoln Laboratory, Massachusetts Institute of Technology
Lexington, Massachusetts 02173

SUMMARY

A tunable high-resolution two-stage heterodyne radiometer has been developed for the purpose of investigating the intensity and lineshape of the 752.033-GHz ($2_{11} \leftarrow 2_{02}$) rotational transition of water vapor. Single-sideband system noise temperatures of approximately 45,000 K have been obtained using a sensitive GaAs Schottky diode as the first-stage mixer. First local-oscillator power was supplied by a CO_2 -laser pumped formic acid laser (761.61 GHz), generating an X-band IF signal with theoretical line center at 9.5744 GHz. Second local-oscillator power was provided by a tunable C-band source. A resolution capability of 1 MHz was achieved by means of a 3-GHz waveguide cavity filter with only 9-dB insertion loss.

In absorption measurements of the H_2O taken from a laboratory simulation of a high-altitude rocket plume, the center frequency of the 752-GHz line was determined to within 1 MHz of the reported value. A rotational temperature ~ 75 K, a linewidth ~ 5 MHz and a Doppler shift ~ 3 MHz (from a 45-degree rotation of the flow direction) were measured with the line-of-sight intersecting the simulated-plume axis at a distance downstream of 30 nozzle diameters. These absorption data were obtained against continuum background radiation sources at temperatures of 1175 and 300 K.

INTRODUCTION

Recent advances in mixer technology at submillimeter-wavelengths have made possible the development of heterodyne receivers with single-sideband system temperatures below 20,000 K (ref. 1). As part of a Lincoln Laboratory program to investigate the submillimeter-wave radiative properties of H_2O molecules in high-altitude rocket plumes (refs. 1,2), a heterodyne receiver similar to that reported earlier (ref. 1) has been adapted to measure the rotational temperature of the 752.033 GHz (ref. 3) transition ($2_{11} \leftarrow 2_{02}$) of H_2O (a prominent plume constituent) under simulated conditions of high altitude (~ 150 km). At altitudes where collisions with ambient molecules are sufficiently infrequent, adiabatic cooling through rapid volume expansion of H_2O vapor emitted into the

* Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91103

vacuum from a nozzle can produce rotational temperatures of selected transitions well below 100 K (refs. 2,4). In work reported recently, Bulatov et al. (ref. 5) performed active measurements with a spectrometer on four submillimeter-wave H_2O lines (including the $2_{11} \leftarrow 2_{02}$ transition) in a planar stream under conditions of relatively high ambient pressure through a point almost at the exit plane. This paper describes the experimental details and some initial results of passive measurements of H_2O emitted from a simulated rocket into a high vacuum at a point farther downstream (ref. 6).

THEORETICAL CONSIDERATIONS

Based on the results of a theoretical analysis (refs. 2,4), it has been concluded that the combination of high optical depth and low rotational temperature in a water vapor vacuum expansion would most likely occur for the low-lying rotational transitions, particularly the $1_{10} \leftarrow 1_{01}$ at ~ 557 GHz and the $2_{11} \leftarrow 2_{02}$ at ~ 752 GHz (fig. 1). At these spectral-line frequencies, temperatures far below room temperature (~ 50 K) are expected to be reached within a short distance from the nozzle through adiabatic cooling by volume expansion.

With the conditions of low pressure and temperature anticipated in these experiments, pressure broadening or self-broadening of the spectral line is not possible in the downstream regime of the plume. Thermal Doppler effects from random molecular motion would normally be the dominant broadening mechanism, and would be determined by

$$\Delta\nu_{\text{thermal}} = 2 \left(\frac{2kT \ln 2}{mc} \right)^{1/2} \nu, \quad (1)$$

where m is the mass of the H_2O molecule, k is Boltzmann's constant, and c is the velocity of light. Since $\Delta\nu_{\text{thermal}} = 2.2$ MHz at $\nu = 752$ GHz and $T = 300$ K, however, it can be argued that the principal source of line broadening in the expanding plume is more likely to be the molecular flow velocities.

Because the molecular flow velocity v_{flow} of the plume can reach several km/sec, its Doppler effects are expected not only to cause line-center frequency shifts $\Delta\nu_{\text{shift}}$ which depend on the direction of the flow relative to the observer, but also to control the linewidth $\Delta\nu_{\text{width}}$, which is a function of v_{flow} and the plume divergence angle. In figure 2, an elementary model depicting these Doppler effects is sketched. With the flow directed at an angle θ to the observer, the frequency shift may be estimated by the relation

$$\Delta\nu_{\text{shift}} \sim \frac{v_{\text{flow}}}{c} \nu \sin\theta. \quad (2)$$

For a plume divergence angle α (defining the optically thick part of the plume), the Doppler linewidth may be estimated by

$$\Delta\nu_{\text{width}} \sim \frac{2v_{\text{flow}}}{c} \nu \sin(\alpha/2). \quad (3)$$

With $\alpha = \pi/2$, and $v_{\text{flow}} = 2$ km/sec, for example, $\Delta\nu_{\text{width}} \sim 7$ MHz at $\nu = 752$ GHz, a linewidth significantly greater than even the room temperature value of the thermal Doppler width estimated above. For a cold plume temperature (~ 75 K), the $\Delta\nu_{\text{thermal}}$ is reduced to 1.1 MHz.

EXPERIMENTAL TECHNIQUES

To measure the rotational temperature of the 752-GHz line, a high-resolution heterodyne radiometer was employed. As shown in figure 3, a formic acid (HCOOH) FIR laser, pumped by a 40-W CO₂ laser at 10.6 μm was used to provide local oscillator power for the first heterodyne stage of the radiometer. The strong cw line at 761.6 GHz is capable of delivering 30 mW and is within 10 GHz of the H₂O rotational line of interest. First-stage mixing was accomplished by means of a GaAs Schottky diode (ref. 1). After intermediate frequency (IF) amplification with a low-noise X-band FET amplifier, the signal was mixed down to S-band by means of a C-band sweep generator to provide an IF for which a 0.9-MHz bandwidth, 3-GHz cavity waveguide filter with only 9-dB insertion loss was utilized (fig. 4). Because the linewidths of thermally broadened rotational lines were on the order of only a few MHz, the importance of a filter capable of providing resolution less than 1 MHz is crucial in any application where molecular densities and optical depths are low enough to produce linewidths that approach the thermal Doppler limit. Another feature of this two-stage receiver was the continuously fine tunability afforded by the second local oscillator.

To make passive measurements of narrow-line spectra with a heterodyne receiver, it is important to remember that, while two sidebands are generated by the mixing process, only one of them will contain the desired information. In figure 5, it is explained how the two sidebands combine in the first IF stage. In the case depicted, a room temperature absorption line in the lower sideband is observed against a hot blackbody continuum, with the baseline for each sideband supplied by the room temperature chopper blade reference. Since the upper sideband is not filtered out, the baseline of the lower sideband must be identified as the top of the upper sideband contribution when the two sidebands are combined.

A convenient feature of this particular experiment lies in the fact that collision broadening of the atmospheric H₂O line is large enough to filter out the entire lower sideband when steps to remove this moisture from the optical path are not taken. As a result, the upper sideband signal contribution can be isolated and recorded initially, and then the volume surrounding optical path of the signal purged in a dry N₂ enclosure (fig. 4) to permit the passage of the lower sideband containing the absorption line. Once this is accomplished,

the resultant signal is doubled (except over the narrow band of the H_2O absorption line) and the measurements can proceed, using the top of the upper sideband as the 300-K reference. By measuring the signal from a liquid-nitrogen-cooled (77 K) blackbody source, the output voltage from the phase-sensitive detector (PSD) can be calibrated directly in terms of temperature. With this calibration source, a single-sideband system noise temperature of 45,000 K was determined.

For the radiometric observations of simulated rocket plumes, a large high-vacuum chamber (~ 10 m long by 3 m in diameter) with a liquid-nitrogen cryopanelled lining was used to create the conditions of high altitude. Along the axis of the chamber, a laboratory-scale jet with a sonic nozzle was used to generate a controlled flow of adiabatically expanding H_2O molecules, as shown in figure 6. To provide a background of temperature greater than that of the reference, in this case a 300-K chopper blade, a mercury-arc lamp was installed at a port opposite to the Z-cut quartz exit window. By defocussing the collimating optics (fig. 7), a real image of the source may be established at the first-stage mixer, with the ray paths parallel to the beam from the FIR local oscillator. With this background, a signal-to-rms noise of ~ 20 was attainable for an integration time of 4 seconds.

EXPERIMENTAL RESULTS

With the chamber evacuated to a roughing pressure of 10^{-2} Torr and the arc lamp turned on, an optically thick absorption of the $2_{11} \leftarrow 2_{02}$ transition was measured with center at 9.5737 ± 0.0010 GHz in the first IF band (table I). In figure 8, the trace of the Lorentzian-shaped line is presented, with a scale determined by calibrating the radiometer by means of a 77 K blackbody (a sheet of absorbing material cooled by liquid nitrogen). The mercury-arc lamp temperature was measured at 1000 K, and raised to 1175 K after correcting for the effect of the 300-K quartz window with emissivity ~ 0.20 . The postdetection integration time for these measurements was 4 seconds. If the frequency of the formic acid LO line is taken as 761.6076 ± 0.0005 GHz (ref. 7), and the H_2O line frequency as 752.0332 ± 0.0005 GHz (ref. 3), the expected first IF is 9.5744 ± 0.0010 GHz, less than 1 MHz from the measured value. The discrepancy between these two values can be accounted for by the accumulated measurement errors involved in the present experiment and the referenced work. The minimum linewidth observed was approximately 2.5 MHz, in good agreement with the calculated thermal (Doppler) width of 2.2 MHz for H_2O at 300 K.

With the chamber walls cooled to 77 K and the pressure reduced to $\sim 10^{-6}$ Torr, the rotational temperature of the H_2O line in the plume of the steam jet was investigated. For a water flow rate of 1 g/sec and a stagnation temperature of 900 K, the theoretical prediction (ref. 4) of the signal brightness or antenna temperature T_A as a function of distance downstream (in terms of nozzle diameters R) along the plume axis is given by the curve in figure 9. In figure 10, the absorption line trace for $R = 30$ is presented. The temperature at the line center was 75 K (after applying the quartz window emissivity correction to the measured value of 120 K), and the linewidth was 5.5 MHz.

In figure 11, the results of a second experiment are presented. To test for the occurrence of a Doppler shift, the steam jet was rotated away from the radiometer through an angle $\theta = \pi/4$ and the absorption measurement repeated. In this case, the line center frequency was red-shifted by 3 MHz, although its width decreased to 4.2 MHz because of the likelihood that a portion of the plume was striking the cold wall before passing through the radiometer beam. Finally, the absorption was observed with a 300-K background (i.e., with the lamp turned off) and the temperature and linewidth were in agreement with the data taken for the hot background.

The results of frequency, linewidth and Doppler shift measurements are listed in table I. From equation (3), an effective angle $\alpha = 80^\circ$ and a flow velocity $v_{\text{flow}} = 1.7$ km/sec (ref. 8) would provide a linewidth of 5.5 MHz. With the plume directed away from the radiometer ($\theta = \pi/4$), and with wall effects taken into account, the calculated value of Δv_{width} for the same plume conditions was 3.8 MHz, in reasonable agreement with the measured value of 4.2 MHz mentioned above. The calculated estimate of Δv_{shift} from equation (2) for these plume conditions was 3.0 MHz, in agreement with the measured value.

CONCLUSIONS

With the aid of a high-resolution two-stage heterodyne radiometer, spectral absorption measurements of the 752.033 GHz line of H_2O vapor were carried out using a blackbody continuum as a background radiation source for investigating the absorptive properties of the H_2O content of simulated high altitude rocket plumes. The receiver had a resolution capability of 1 MHz and was tunable through the use of a C-band microwave sweep generator as a second-stage local-oscillator source. With thermally Doppler broadened H_2O at room temperature, the line center frequency was determined within 1 MHz of the reported value, and the linewidth was in close agreement with the calculated value.

To simulate the physical situation of a high altitude rocket in a laboratory environment, a small steam jet was operated inside a large high vacuum chamber, with the H_2O jet plume traversing the radiometer line-of-sight. These experiments verified that this rotational line is optically thick, with excitation temperatures below 100 K, in the downstream part of the plume, as predicted by theoretical modeling. From the measured traces of the absorption line profile, it was concluded that Doppler effects from plume molecular flow velocity components along the radiometer line-of-sight were the principal sources of the measured linewidth.

A Doppler shift in the line center frequency which occurred when the plume direction was rotated relative to the line-of-sight was also judged to be caused by the change in direction of the molecular flow velocity component along the plume axis, as verified by the agreement between the experimental result and theoretical prediction.

The authors are indebted to J. A. Weiss for plume signature calculations, to B. J. Clifton, W. D. Fitzgerald and P. E. Tannenwald for development of the

technology required to carry out these measurements, to C. D. Parker for assistance in the operation of the heterodyne receiver, and to W. E. Courtney for providing the narrow-band filter. This work was sponsored by the Department of the Army.

REFERENCES

1. Fetterman, H. R.; Tannenwald, P. E.; Clifton, B. J.; Parker, C.D.; Fitzgerald, W. D.; and Erickson, N. R.: Far-IR Heterodyne Radiometer Measurements with Quasioptical Schottky Diode Mixers. Appl. Phys. Lett. 33, 151 (1978).
2. Dionne, G. F.; Litvak, M. M.; Weiss, J. A.; and Fitzgerald, W. D.: Detectability of Cold Rocket Plumes. Technical Note 1979-48, Lincoln Laboratory, M.I.T., (11 October 1979). AD-A081858.
3. DeLucia, F. C.; Helminger, P.; Cook, R. L.; and Gordy, W.: Submillimeter Microwave Spectrum of $H_2^{16}O$. Phys. Rev. A5, 487 (1972).
4. Litvak, M. M.; Weiss, J. A.; and Dionne, G. F.: Submillimeter-Wave Properties of Thermospheric Rocket Plumes. Digest of the Fourth International Conference on Infrared and Millimeter Waves and their Applications, p. 139, (1979).
5. Bulatov, E. D.; Vinogradov, E. A.; Konuykhov, V. K.; Prokhorov, A. M.; Sisakyan, I. N.; Tikhonov, V. I.; and Shiskov, P. O.: Rotational Non-equilibrium of $H_2^{16}O$ Molecules in a Supersonic Rarefied Water Vapor Stream. Sov. Phys. JETP 49, 274 (1979).
6. Dionne, G. F.; Fitzgerald, J. F.; Chang, T-S.; and Litvak, M. M.: Radiometric Observations of the 752.033-GHz Absorption Line in Simulated Rocket Plumes. Digest of the Fourth International Conference on Infrared and Millimeter Waves and their Applications, p. 133 (1979).
7. Deldalle, A.; Dangoisse, D.; Springard, J. P.; and Bellet, J.: Accurate Measurements of cw Optically Pumped FIR Laser Lines of Formic Acid Molecule and its Isotopic Species $H^{13}COOH$, $HCOOD$ and $DCOOD$. Opt. Comm. 22, 333, (1977).
8. Sutton, E. A.: Design Considerations for a Laboratory Jet Simulating Rocket Exhausts. Concord Sciences Corporation Report CSC RN-4 (1977).

TABLE I.-SUMMARY OF EXPERIMENTAL RESULTS FOR $R = 30$, $v_{\text{flow}} = 1.7$ km/sec, and $\alpha = 80^\circ$

	<u>Measured</u>	<u>Calculated</u>
First IF (GHz)	9.5737 ± 0.0010	9.5744 ± 0.0010
Plume Temperature (K)	75	80
Δv_{width} with $\theta = 0$ (MHz)	5.5	5.5
Δv_{width} with $\theta = \pi/4$ (MHz)	4.0	3.8
Δv_{width} for $\theta = \pi/4$ (MHz)	3.0	3.0

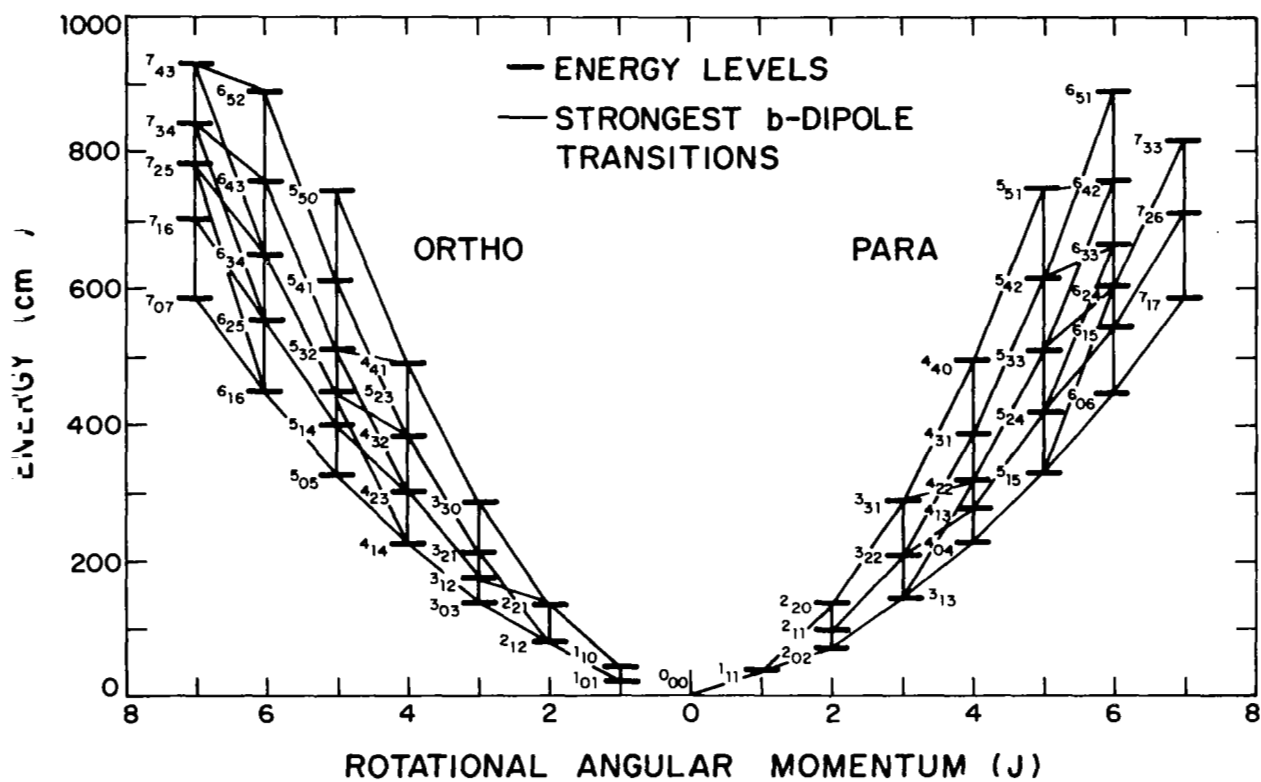


Figure 1.- Water molecule rotational energy levels.

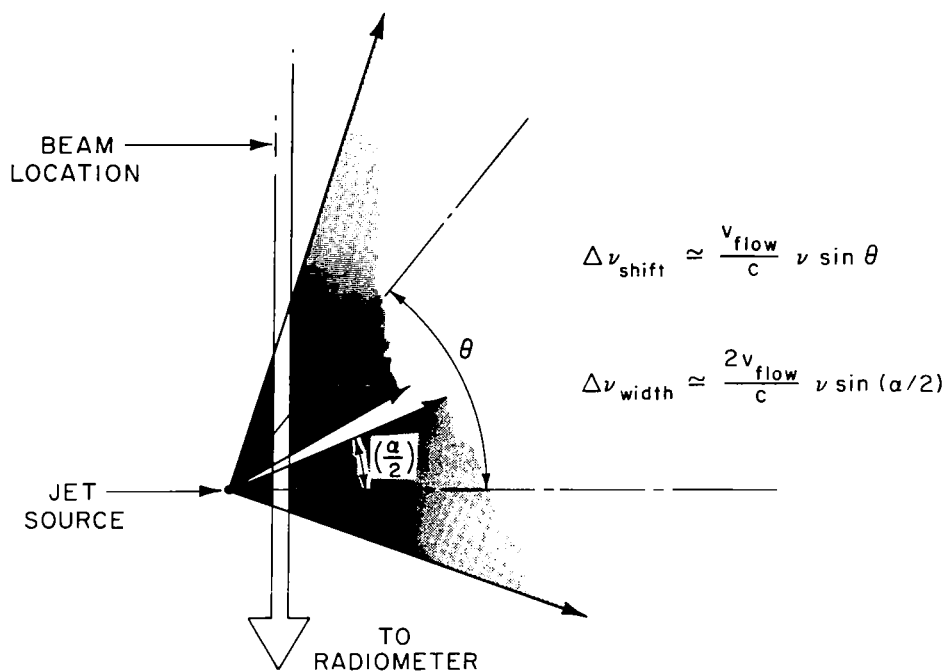


Figure 2.- Elementary model section of plume expansion into a vacuum, defining the plume angle α and the plume direction angle θ relative to the radiometer line-of-sight.

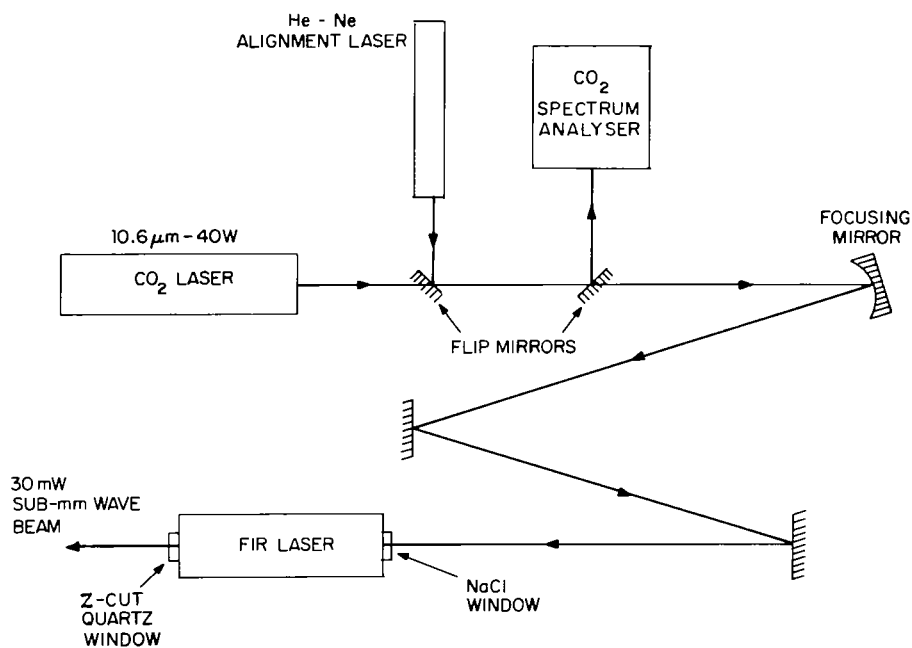


Figure 3.- Optical arrangement for submillimeter-wave (FIR) laser with CO₂ laser pump source.

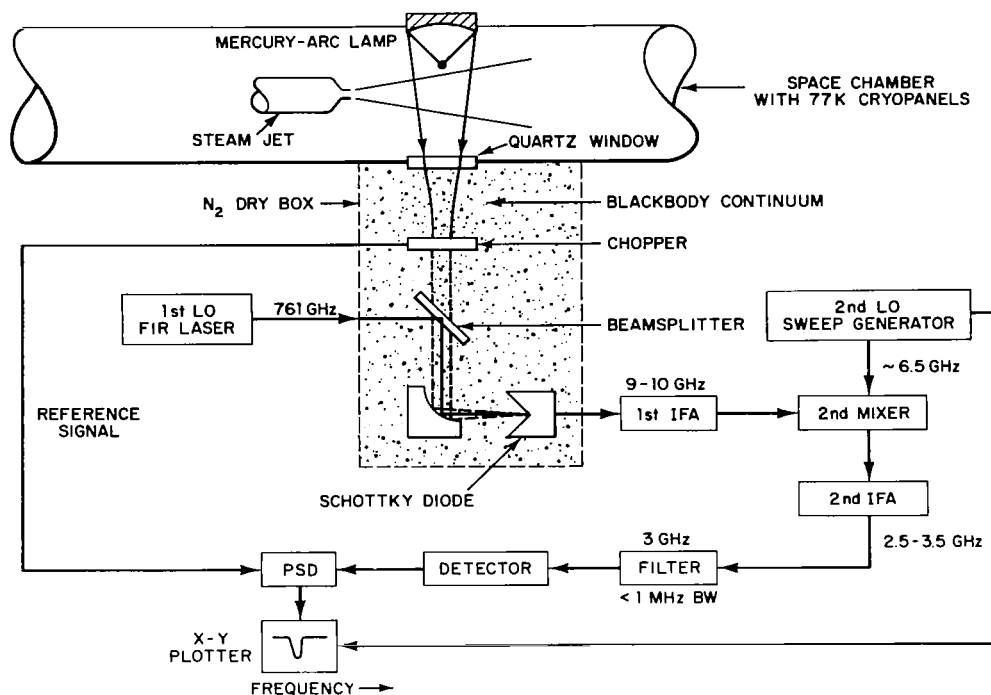


Figure 4.- Block diagram of two-stage heterodyne radiometer for operation at 752 GHz, with high-vacuum space chamber and water vapor jet.

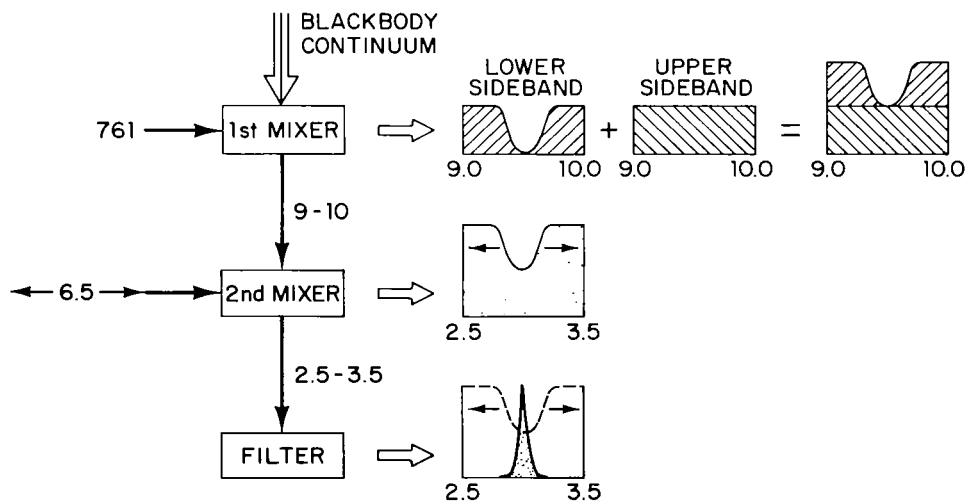


Figure 5.- Single-sideband detection of a narrow absorption line in the presence of two sidebands, using a hot continuum background source.

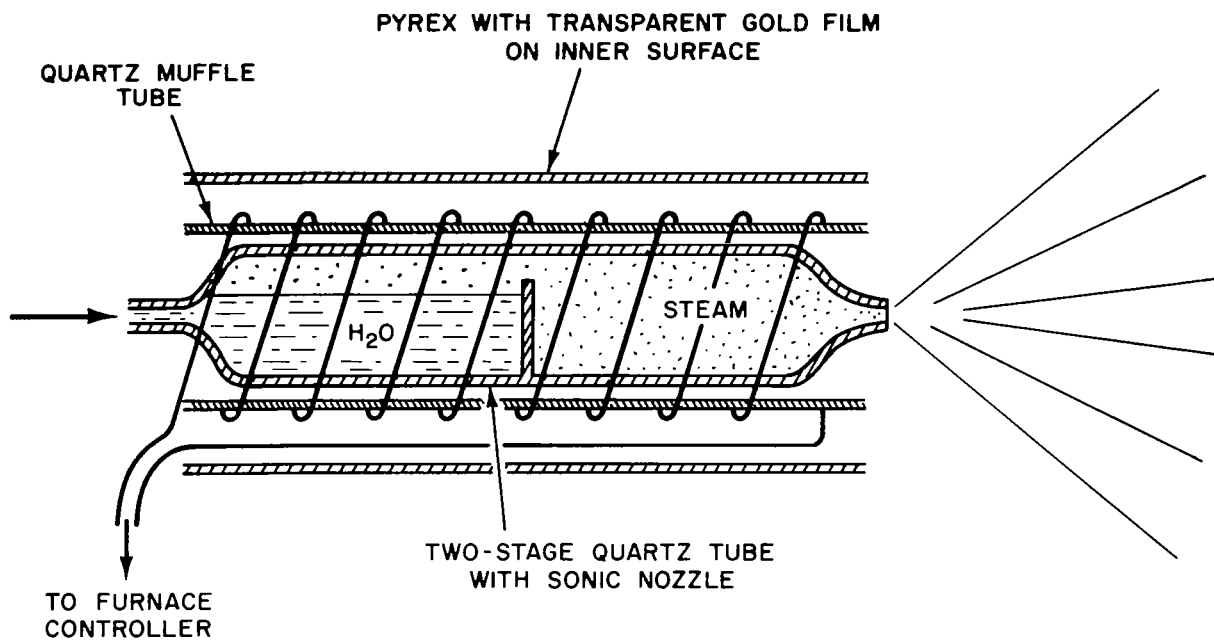


Figure 6.- Laboratory-scale water vapor jet with sonic nozzle.

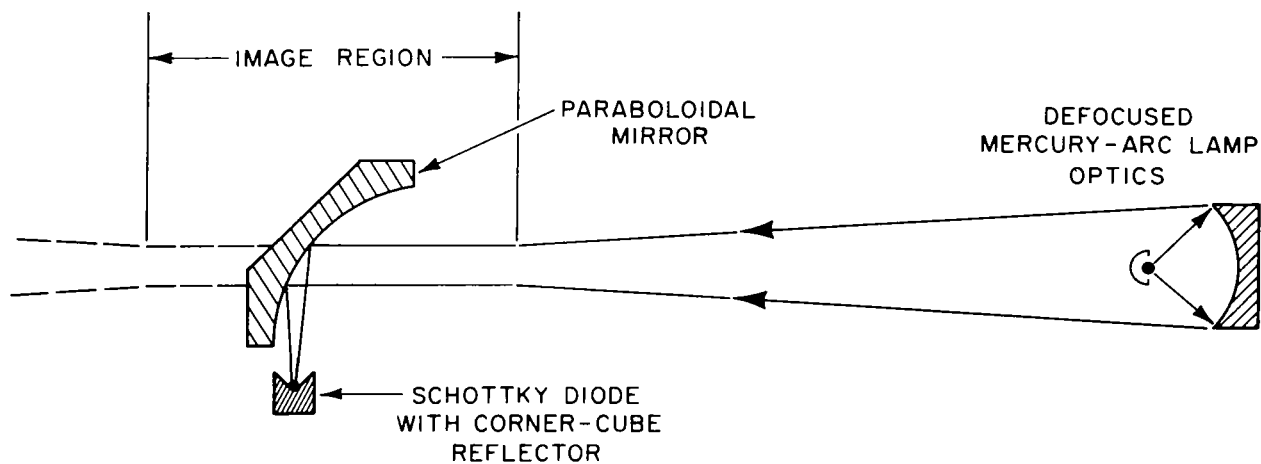


Figure 7.- Optical arrangement for the mercury-arc lamp, paraboloidal mirror and Schottky-barrier diode mixer.

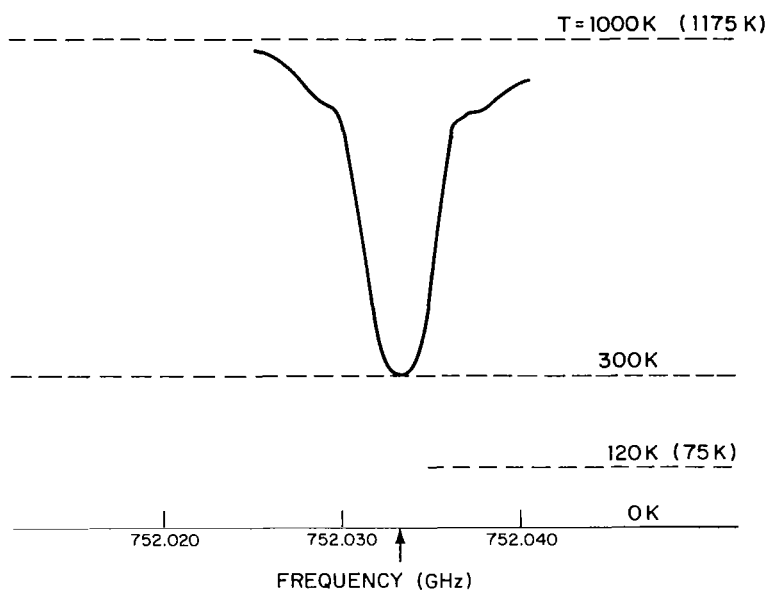


Figure 8.- Experimental trace of thermally broadened H_2O line at 752.033 GHz with vacuum chamber walls at 300 K.

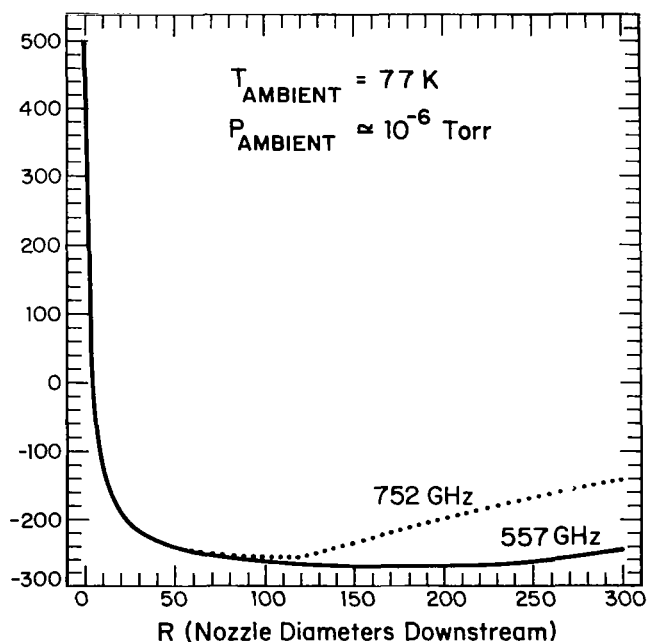


Figure 9.- Theoretical prediction of jet plume brightness temperature in laboratory rocket simulation experiment as a function of R (nozzle diameters downstream).

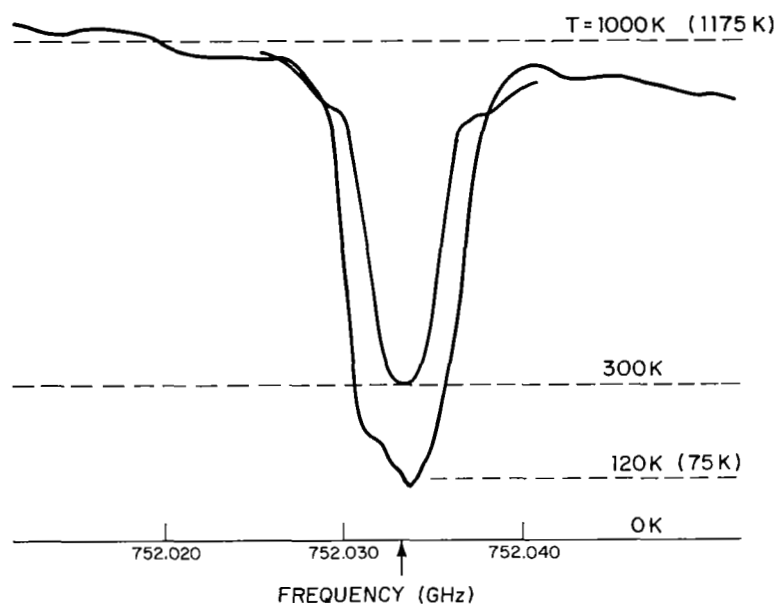


Figure 10.- Experimental trace of plume H_2O line at 752.033 GHz with vacuum chamber walls at 77 K and for $R = 30$ and $\theta = 0$.

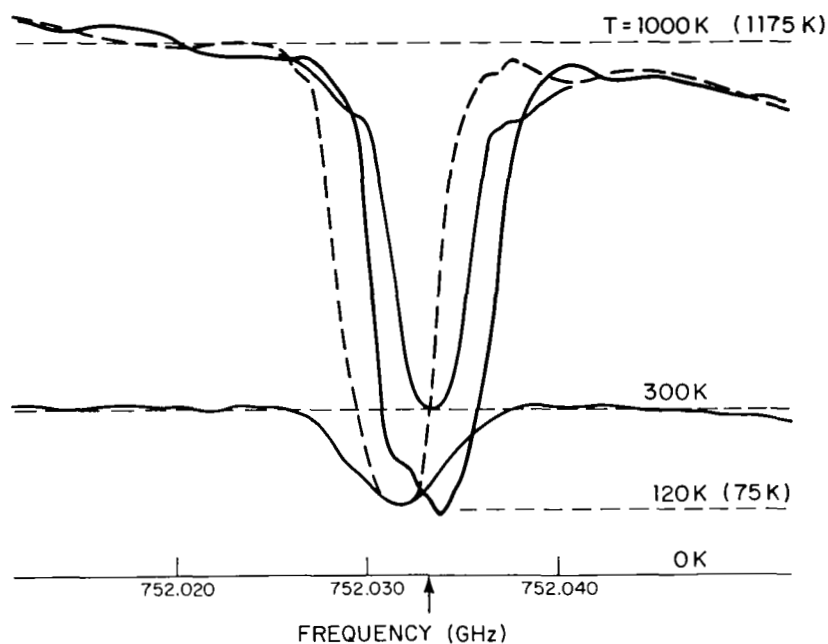


Figure 11.- Experimental trace of H_2O line at 752.033 GHz with vacuum chamber walls at 77 K for the case of $R = 30$ and $\theta = \pi/4$, showing the Doppler shift. Data are presented for both 1175 K and 300 K continuum background cases.

1. Report No. NASA CP-2138, Part 1		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle HETERODYNE SYSTEMS AND TECHNOLOGY				5. Report Date August 1980	
				6. Performing Organization Code	
7. Author(s)				8. Performing Organization Report No. L-13849	
9. Performing Organization Name and Address NASA Langley Research Center Hampton, VA 23665				10. Work Unit No. 506-61-53-04	
				11. Contract or Grant No.	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546				13. Type of Report and Period Covered Conference Publication	
				14. Sponsoring Agency Code	
15. Supplementary Notes					
16. Abstract <p>This report is a compilation of papers presented at the International Conference on Heterodyne Systems and Technology held March 25-27 at Williamsburg, Virginia. It includes papers that cover various aspects of optical heterodyning including active and passive systems, local oscillators, tunable diode lasers, detectors, and related systems.</p>					
17. Key Words (Suggested by Author(s)) Heterodyne Optical Laser Photomixer			18. Distribution Statement Unclassified - Unlimited Subject Category 36		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 270	22. Price* A12		